Claimed effects, outcome variables and methods of measurement for health claims on foods related to the gastrointestinal tract proposed under Regulation (EC) 1924/2006.

Beatrice Biasini[1#], Laura Marchi[1#], Donato Angelino[1], Giorgio Bedogni[2], Ivana Zavaroni[3,4], Carlo Pruneti[5], Daniela Galli[6], Prisco Mirandola[6], Marco Vitale[6], Alessandra Dei Cas[3,4], Riccardo C. Bonadonna[3,4], Giovanni Passeri[7], Marco Ventura[8], Daniele Del Rio[1], Daniela Martini[1*].

[1] The Laboratory of Phytochemicals in Physiology, Department of Food and Drugs, University of Parma, Parma, Italy.

[2] Clinical Epidemiology Unit, Liver Research Center, Basovizza, Trieste, Italy.

[3] Division of Endocrinology, Department of Medicine and Surgery, University of Parma, Parma, Italy.

[4] Azienda Ospedaliera Universitaria of Parma, Parma, Italy.

[5] Department of Medicine and Surgery, Clinical Psychology Unit, University of Parma, Parma, Italy.

[6] Sport and Exercise Medicine Centre (SEM), Department of Medicine and Surgery, University of Parma, Parma, Italy.

[7] Unit of Andrology, Metabolic Bone Diseases and Endocrinology, Department of Medicine and Surgery, University of Parma, Parma, Italy.

[8] Laboratory of Probiogenomics, Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Parma, Italy.

*Corresponding author: Dr. Daniela Martini. Department of Food and Drugs, University of Parma, Parma, Italy. Telephone: +390521903913. E-mail: daniela.martini@unipr.it

[#]Authors equally contributed to the study.

ABSTRACT

Most of the requests of authorization to the use of health claims pursuant to Regulation EC 1924/2006 related to the gastrointestinal (GI) tract have received a negative opinion by the European Food Safety (EFSA), mainly because of an insufficient substantiation of the claimed effect (CE).

The present manuscript refers to the collection, collation and critical analysis of outcome variables (OVs) and methods of measurement (MMs) related to the GI tract compliant with Regulation 1924/2006.

The critical evaluation of OVs and MMs was based on the literature review, with the final aim of defining their appropriateness in the context of a specific CE. The results obtained are relevant for the choice of the best OVs and MMs to be used in randomized controlled trials aimed to substantiate the claims on the GI tract. Moreover, results can be used by EFSA for updating the guidance for the scientific requirements of such health claims.

Keywords: health claim; claimed effect; outcome variable; method of measurement; gastrointestinal tract.

List of abbreviations:

BMC: Bone Mineral Content; **BMD**: Bone Mineral Density; **BSQ**: Bowel Symptom Questionnaire; **CE**: Claimed Effect; **CTT**: Colonic Transit Time; **D-IBS**: Diarrhoea-predominant IBS; **DXA**: Dual-Energy X- Ray Absorptiometry; **EFSA**: European Food Safety Authority; **FDDQOL**: Functional Digestive Disorders Quality Of Life; **Gastro-Q**: Gastro-Questionnaire; **GI**: Gastrointestinal; **GSRS**: Gastrointestinal Symptom Rating Scale; **IBS**: Irritable Bowel Syndrome ; **IBS-SS**: Irritable Bowel Syndrome-Symptom Severity Scale; **ITS**: Internal Transcribed Space; **MM**: Method of Measurement; **MPQ**: McGill Pain Questionnaire; **MRI**: Magnetic Resonance Imaging; **NAA**: Neutron-Activation Analysis; **NP**: Nausea Profile; **OCTT**: Orocecal Transit Time; OV: Outcome Variable; **PAC**: Patient Assessment of Constipation; **PAC-QOL**: Patient Assessment of Constipation- Quality Of Life; **PAC-SYM**: Patient Assessment of Constipation- Symptom; **PRO**: Patient Reported Outcome; **QCT**: Quantitative computed tomography; **QOL**: Quality Of Life; **ROM**: Radio-Opaque Markers; **SF-36**: Short Form-36; **SF-MPQ**: Short Form version of MPQ; **SGA**: Subjective Global Assessment; **SIBO**: Small Intestinal Bacterial Overgrowth; **SPA**: Single Photon Absorptiometry; **SST**: Single Stool Transit; **VAS**: Visual Analogue Scale; **WGTT**: Whole-gut transit time; **WHO**: World Health Organization.

Index

3.1.3.3. STOOL WEIGHT/VOLUME/SIZE
3.1.3.3.1. DIRECT ASSESSMENT BY THE INVESTIGATORS
3.1.3.4. INTESTINAL TRANSIT TIME
3.1.3.4.1. ROM TECHNIQUE
3.1.3.4.2. SST WITH COLOURED PLASTIC PELLETS
3.1.3.4.3. HYDROGEN BREATH TEST
3.1.3.5. DIARRHOEA
3.1.3.5.1. QUESTIONNAIRE
3.1.3.6. BLOATING
3.1.3.6.1. QUESTIONNAIRE
3.1.3.7. BORBORYGMI
3.1.3.7.1. QUESTIONNAIRE
3.1.3.8. FLATULENCE
3.1.3.8.1. QUESTIONNAIRE
3.1.3.9. ABDOMINAL PAIN/CRAMPS
3.1.3.9.1. VISUAL ANALAOGUE SCALE
3.1.3.9.2. QUESTIONNAIRE
3.1.3.10. FAECAL BACTERIAL MASS
3.1.3.10.1. GRAVIMETRIC PROCEDURE
3.1.3.11. COMPOSITION OF THE GUT MICROBIOTA/ BIFIDOBACTERIAL POPULATION
3.1.3.11.1. 16S rRNA MICROBIAL PROFILING
3.1.3.11.2. BIFIDOBACTERIA ITS PROFILING
3.1.3.12. QUALITY OF LIFE
3.1.3.12.1. FUNCTIONAL DIGESTIVE DISORDERS QUALITY OF LIFE QUESTIONNAIRE
3.1.3.13. CONSTIPATION
3.1.3.13.1. PATIENT ASSESSMENT OF CONSTIPATION
3.1.3.14. SENSATION OF COMPLETE/INCOMPLETE EVACUATION
3.1.3.14.1. QUESTIONNAIRE
3.1.4. IMPROVING IRON ABSORPTION
3.1.4.1. NON- HAEM IRON ABSORPTION
3.1.4.1.1. DOUBLE ISOTOPE TECHNIQUE
3.1.4.1.2. WHOLE BODY COUNTING
3.1.5. IMPROVING OF LACTOSE DIGESTION
3.1.5.1. HYDROGEN BREATH CONCENTRATION
3.1.5.1.1. HYDROGEN BREATH TEST
3.1.5.2. NAUSEA
3.1.5.2.1. QUESTIONNAIRE
3.1.5.3. DIARRHOEA
3.1.5.3.1. QUESTIONNAIRE
3.1.5.4. ABDOMINAL PAIN/CRAMPS
3.1.5.4.1. VISUAL ANALOGUE SCALE
3.1.5.4.2. QUESTIONNAIRE
3.1.5.5. BLOATING
3.1.5.5.1. QUESTIONNAIRE
3.1.5.6. FLATULENCE
3.1.5.6.1. QUESTIONNIARE
3.2. <u>CLAIMS REFERRING TO CHILDREN'S DEVELOPMENT AND HEALTH</u>
3.2.1. REDUCTION OF GI DISCOMFORT
3.2.1.1. CRYING TIME AND FREQUENCY
3.2.1.1.1. PARENTS' DIARY
3.2.1.2. ABDOMINAL DISTENSION
3.2.1.2.1. PARENTS' DIARY
3.2.1.2.2. PARENTAL INTERVIEW

1. INTRODUCTION

Gastrointestinal (GI) disorders may derive from several different diseases or situations, and are characterized by a wide spectrum of symptoms. The most referred symptoms to the family physicians are bloating and abdominal pain, and it has been established that each person in the life had experienced at least an episode of both (Iovino et al. 2014, Knowles and Aziz 2009). These are very general symptoms, usually followed by an alteration of the stool consistency and frequency of evacuation, ranging from constipation to diarrhoea (Viniol et al. 2014). Although they can turn into several pathologies, *i.e.* faecal impaction, incontinence or bowel perforations, the simply ascertaining of the presence of these symptoms may impact the quality of life of the individual, affecting both the mental and the behavioural state in children and in adults (Belsey et al. 2010, Borgaonkar and Irvine 2000). For these reasons, validated questionnaires have been developed in order to qualify and quantify the discomforts and help physicians in the formulation of diagnosis (Belsey et al. 2010, Borgaonkar and Irvine 2000, Wald and Sigurdsson 2011, da Fonseca 2015). The impact of lifestyle behaviours on gut function has been widely studied: for instance, it has been claimed that smoking negatively affects the correct functionality of the GI tract (Li et al. 2014), while a constant physical activity, in a relatively low intensity, has a protective effect on gut functions (Peters et al. 2001). Furthermore, diet seems to have a strong impact on gut, as foods or dietary patterns may act both in a negative or positive way on its function. Data reported by The first National Health and Nutrition Examination Survey (NHANES-I) on self-reporting constipation and dietary interviews on more than fifteen thousand volunteers evidenced that the consumers of fruit and vegetables, milk and poultry had fewer episodes of constipation, higher among strong consumers of tea and coffee (Sandler et al. 1990). Prebiotics (*i.e.* fibre) and probiotics have been the most studied food/food components for their role on gut functions, but to date there is still a debate on their effects. Concerning dietary fibre, contrasting results have been found when dietary fibre intake has been correlated with bowel movements or constipation (Murakami et al. 2006, Sanjoaquin et al. 2004). Similarly, reviews and meta-analyses evidenced that probiotics have a beneficial role on some markers of gut function, *i.e.*

242 stool consistency, but interpretation of the data are still debated due to their high heterogeneity and
243 risk of bias (Dimidi et al. 2014, Martinez-Martinez et al. 2017).
244 A variety of foods and food components, including dietary fibre, probiotic bacteria and yeasts, have
245 been the object of applications for authorization of health claims pursuant to Regulation (EC)
246 1924/2006. Most of them have received a negative opinion by the European Food Safety Authority
247 (EFSA) due to a variety of reasons ranging from the non-exhaustive characterization of the
248 food/food constituent to the inappropriate formulation or the insufficient substantiation of the
249 claimed effect. For instance, many negative opinions were due to methodological limitations of the
250 studies provided by applicants, including the choice of not appropriate outcome variables (OVs)
251 and/or methods of measurement (MMs).
252 In this scenario, a project focusing on the appropriateness of the OVs and MMs selected by the
253 applicants has been developed, as described in previous manuscripts (Martini et al. 2018a, Martini
254 et al. 2017a, Martini et al. 2018b, Martini et al. 2017b), with the aim to improve the quality of
255 applications provided to EFSA. The present manuscript refers to the collection, collation and critical
256 analysis of OVs and MMs related to gastrointestinal tract functions, excluding immune function,
257 compliant with the Regulation 1924/2006.
258 **2. MATERIALS AND METHODS: SEARCH STRATEGY**
259 OVs and MMS were collected from the relative Guidance document (EFSA 2016) and from the
260 requests for authorization of health claims under Article 13.5 and 14 of the Regulation (EC)
261 1924/2006 related to GI tract functions (http://ec.europa.eu/nuhclaims/). As described by Martini *et*
262 *al.* (Martini et al. 2017b), the OVs and MMs were included only if the food/food constituent(s) was
263 sufficiently characterized and the claimed effect was considered to be beneficial. Following this
264 decision tree, 5 claimed effects related to the GI tract, with the exclusion of immune functions, with
265 38 OVs were evaluated under Article 13.5. Moreover, 4 claimed effects with 15 OVs referred to
266 children development were selected under Article 14. For each OV, a database of references was
267 created on PubMed and was used for the critical analysis of the OVs and the MMs. Each OV and

related MM was ranked in one of the following categories: (i) appropriate; (ii) appropriate only/better if in combination with other OV or MM; (iii) not appropriate *per se*; (iv) not appropriate in relation to the specific claimed effect proposed by the applicant(s), (v) not appropriate alone, but useful as supportive evidence for the scientific substantiation of the claimed effect.

3. CRITICAL ANALYSIS OF OUTCOME VARIABLES AND METHODS OF MEASUREMENT

3.1 FUNCTION HEALTH CLAIMS

3.1.1. REDUCTION OF GI DISCOMFORT

3.1.1.1 SUBJECTIVE GLOBAL ASSESSMENT OF SYMPTOMS

Subjective Global Assessment (SGA) of symptoms is a tool allowing the evaluation of several GI symptoms integrating the results obtained for each symptom in a single parameter. The choice of symptoms to be included in a SGA depends on the particular GI disorder or health claim to be evaluated. In fact, the parameter obtained with a SGA includes measures of change for each of the symptoms which are part of the entry criteria. In the context of GI disorders, SGA includes the evaluation of changes in GI discomfort (e.g. bloating, abdominal pain/cramps, straining and borborygmi) and in defecation habits.

To evaluate the appropriateness of SGA of symptoms as OV of reduction of GI discomfort, the literature deriving from database #1 was critically evaluated (Table 1).

Several individual symptoms, which may interact in complex ways, are associated with GI discomfort. Their assessment is not always easy because such symptoms can vary from patient to patient and from time to time, in intensity and duration, and no symptom represents a sufficiently validated parameter to be recommended unequivocally as the primary outcome measure for the substantiation of health claims on the reduction of GI discomfort in general. For these reasons, key symptoms characterising GI discomfort need to be integrated in a single assessment that it is able to represent an overall effect of the intervention of this outcome. Owing the fluctuating nature of GI symptoms, the effect of an intervention should be assessed for extended periods of time (e.g. 4-8

weeks) in order to obtain meaningful results (Irvine et al. 2016, Irvine et al. 2006).

In conclusion, the measurement of SGA of symptoms is an appropriate outcome variable to be used for the substantiation of health claims in the context of reduction of GI discomfort.

3.1.1.1.1. QUESTIONNAIRE

The most important outcomes for evaluation of GI discomfort are the patient's symptoms, such as abdominal pain, bloating, abdominal distention, borborygmus, flatulence, diarrhoea, constipation, bowel urgency, sensation of incomplete evacuation and straining, and patient's defecation habits (e.g. stool frequency, consistency, weight, volume). The intensity, severity and frequency of symptoms can vary from patient to patient and from time to time. In the absence of validated biomarkers allowing objective measures of these symptoms, patient reported outcomes (PRO) are generally accepted (Spiegel et al. 2010). Validated self-administered questionnaires are the recommended method of measurement, because the physician assessment will be less accurate or reliable than the patient's assessment. Diaries and interviews could overcome the problem of recall bias, although their validity should be considered (Irvine et al. 2016, Irvine et al. 2006).

A validated questionnaire for SGA of symptoms must include relevant and representative symptoms of the disorder; moreover, the measure must be reproducible and responsive, and a change in the outcome measures should reflect a real change in general health status. Concerning the severity of the symptoms, the two most used scales are categorical ones (often referred to as Likert scales) and Visual Analogue Scale (VAS). Generally, five or seven-point Likert scales are preferable because they are able to detect small but potentially relevant differences (Muller-Lissner et al. 2003). Most questionnaires assess the severity of symptoms, but some of them take also into consideration frequency and/or duration of symptoms. The choice of particular questionnaire depends on the symptoms or disorder to be monitored, the study group and setting of the study.

Questionnaires frequently used for the assessment of SGA are: GSRS (Gastrointestinal Symptom Rating Scale), IBS-SS (Irritable Bowel Syndrome- Symptom Severity Scale), Gastro-Q (Gastro-Questionnaire), and BSQ (Bowel Symptom Questionnaire).

The GSRS is an interview-based rating scale, easy to apply, consisting of 15 items. It is validated for the assessment of GI symptoms in IBS and peptic ulcer disease. All items are rated in seven steps, of which 0, 1, 2, and 3 are defined by descriptive anchors (0 indicates absence of symptoms and 3 an extreme degree of the symptom). The intensity of symptoms, frequency of attacks, duration of attacks, and their impact on daily living are assessed in the GSRS, when appropriate (Svedlund et al. 1988).

The IBS-SS contains severity scoring questions (related to pain, abdominal pain, abdominal distension, bowel habits and quality of life). Each of the five questions generates a maximum score of 100 using prompted VAS, leading to a total maximum score of 500. It is validated in IBS patients, in which this scoring system produces a meaningful value that is reproducible and sensitive to change (Francis et al. 1997).

The Gastro-Q contains 27 gastrointestinal symptom items drawn from the Rome–II criteria, which are rated by frequency (rated on a 4-point scale) and severity (rated on a 5-point scale), as well as some items to exclude organic diseases. Gastro-Q has been validated in normal participants and in patients with IBS. The Gastro-Questionnaire is a very economic, reliable, and content-valid instrument for the assessment of GI symptoms (Leibbrand et al. 2002).

The self-report BSQ contains 83 items, among which questions on age, sex, marital status, highest level of educational training and employment of the highest income earner in the household (to calculate socioeconomic status). Thirty-six items regard GI symptoms, while 4 are related to health care seeking. The BSQ has been validated in an Australian population-based sample, composed by outpatients, volunteers and random sample of the population. This questionnaire is well accepted, easy to understand, and provides reliable and valid data for assessing GI symptoms (Talley et al. 1995).

In conclusion, validated questionnaires are an appropriate method for the subjective global and individual assessment of GI symptoms. In addition, they are appropriate methods to assess single domains of GI symptoms (bloating, straining, borborygmi, sensation of complete/incomplete

evacuation, abdominal distension, flatulence, need to defecate/bowel urgency, diarrhoea, stool frequency).

3.1.1.2 ABDOMINAL PAIN/CRAMPS

Abdominal pain (also called stomachache) is a pain that occurs between chest and pelvic regions. It can be crampy, achy, dull, intermittent or sharp and may derive from many conditions including infection, presence of abdominal mass, inflammation, obstruction, menstruation, lactose intolerance and intestinal disorders.

To evaluate the appropriateness of abdominal pain/cramps as OV of reduction of GI discomfort, the literature deriving from database #2 was critically evaluated (Table 1).

Abdominal cramping and pain are the central symptoms of IBS, a functional GI disorder characterized by chronic or recurrent abdominal pain or discomfort. The onset of these symptoms reduces the quality of life of affected individuals. Severity is the main recorded characteristic of pain, while less is known about the impact of other pain dimensions, including frequency and duration. Abdominal pain and discomfort are wrongly combined into the same symptom but their distinction is essential for a valid measurement (Spiegel et al. 2010). In fact, abdominal pain often co-exists with one or more symptoms of GI discomfort, such as borborygmi, distension, straining or flatulence. Key symptoms characterising a particular GI disorder, therefore, need to be integrated in a single assessment that it is able to represent an overall effect of the intervention of GI discomfort. Pain is measured separately from discomfort by using a numeric rating scale.

Abdominal pain is also characteristic of lactose maldigestion, although its diagnosis is not solely based on the presence of this unspecific symptom (Jellema et al. 2010).

In conclusion:

- The incidence and severity of abdominal pain/cramps alone are not appropriate to be used as outcome variable for the substantiation of claims in the context of reduction of GI discomfort. A SGA of all symptoms combined should be used instead. Moreover, these outcome variables are not appropriate to be used alone for the substantiation of such claims in children.

- The incidence and severity of abdominal pain/cramps are not appropriate outcome variables for the substantiation of claims in the context of maintenance of normal defecation.
- The incidence and severity of abdominal pain/cramps are not appropriate outcome variables to be used alone for the substantiation of claims in the context of improved lactose digestion, but can be used as supportive evidence for such health claims.

3.1.1.2.1 VISUAL ANALOGUE SCALE

The VAS is a widely used method for the assessment of pain severity and relief. It is reproducible, easy to use, and can be applied to a variety of clinical practices and research. In general, VAS has been developed to measure a parameter that is believed to range across a continuum of values and therefore not directly measurable. Operationally, a VAS is a vertical or horizontal line, 100 mm long, flanked at each end by word descriptors. The patient is asked to rate his current pain perception by drawing a line on a continuous scale from 1 to 10. "1" corresponds to a mild discomfort from time to time, while "10" means the most intense pain. Distance from these two points of the line corresponds to the different degrees of severity. VAS is subjective and useful to assess changes within individual, but less of value for cross-sectional comparisons of different individuals. Validation studies have shown high reliability of VAS in measuring both acute and chronic pain. When a VAS is repeated within a short period of time, 90% of the intra-individual scores usually overlap. Therefore, the repeatability of VAS is good. VAS is also very sensitive to change. From a clinical point of view, a difference of about 13 mm on a VAS represents, on average, a significant change (Gallagher et al. 2001, Williamson and Hoggart 2005).

Based on these considerations, VAS is a solid and appropriate method, for the assessment of severity of abdominal pain/cramps.

3.1.1.2.2 QUESTIONNAIRE

The use of retrospective questionnaires is an acceptable method, provided that the recall interval is limited to the previous 3 months. Questionnaires must be completed before treatment and at follow-up visits. A binary PRO end point, such as "adequate relief," "satisfactory relief," or "considerable

relief", corresponds to a dichotomous responder status (yes/no relief) and represents a primary outcome measure. The patients who give the affirmative response to adequate/satisfactory relief at half of the treatment time, at minimum, are considered as responders. Binary end points are easy to administer and straightforward to interpret, but fail to detect small changes of potential clinical relevance.

One of the most frequently used questionnaire is the McGill Pain Questionnaire (MPQ), a multidimensional pain tool which measures the sensory (what the pain feels like physically), affective (what the pain feels like emotionally), evaluative (overall intensity of the pain experience), and miscellaneous aspects of pain. It is easy to administer and evaluate, as no training is required to score and interpret it. It comprises the Pain Rating Index, and a 1-item, 5-point pain intensity scale (Present Pain Intensity). The Pain Rating Index is composed by 78 pain descriptor items divided into 20 subclasses. Each of them contains 2–6 words referring to 4 major subscales: sensory (subclasses 1–10), affective (subclasses 11–15), evaluative (subclass 16), and miscellaneous (subclasses 17–20). The value (score) is based on 3 main measures: 1) the pain rating index; 2) the number of words chosen; 3) the present pain index based on a 0-5 intensity scale (none (0), excruciating (5)) (Hawker et al. 2011). A higher score on the MPQ indicates the most intense pain. Several studies have been made to validate MPQ and have confirmed the feasibility, reliability, responsiveness and ease of administration of this questionnaire. These studies have been carried out in patients with rheumatoid arthritis (RA) or cancer to evaluate the validity of MPQ in measuring different aspects of pain. However, some patients (older people or illiterate) have difficulty to complete the questionnaire due to the complexity of the vocabulary used. In these cases is necessary supervision during completion of MPQ (Melzack 1975). A short version of MPQ (SF-MPQ) is available for use its application in specific research settings in case of limited time form the patients. (Hawker et al. 2011). In conclusion, questionnaire, e.g. the MPQ, is an appropriate method for assess abdominal pain/cramps.

3.1.1.3 BLOATING

Bloating (or abdominal bloating) is the subjective sensation associated with abdominal distension

(objective sign). Although somehow related, abdominal bloating and distension are two separate symptoms. Bloating affects 10%-30% of the general population and up to 96% of patients with functional gastrointestinal disorders, like functional dyspepsia or IBS, and it is frequently associated with constipation. It is often described by patients as very intrusive, significantly impacting their quality of life. The classification, pathophysiology, clinical significance and treatment of abdominal bloating remain unknown (Houghton 2011, Iovino et al. 2014).

To evaluate the appropriateness of bloating as OV of reduction of GI discomfort, the literature deriving from database #3 was critically evaluated (Table 1).

Bloating is an ambiguous term that can indicate many sensations, like swollen/distended abdomen, full belly, feeling of abdominal pressure or wall tension, or sensation of excess gas; therefore, it can be very subjective (Azpiroz and Malagelada 2005). Bloating is one of the most common and bothersome symptoms for IBS patients (Iovino et al. 2014). Being a subjective symptom, no measurable parameters exist to evaluate the frequency, severity and duration of bloating, especially by a physician. Bloating is also a symptom of carbohydrate malabsorption, especially lactose, but it is not specific and only occurs in about one-third of lactose "malabsorbers" (Azpiroz et al. 2015). Bloating often co-exists with one or more of borborygmi, distension, abdominal pain or flatulence. For that reason, the evaluation of the effect of an intervention on GI discomfort requires the assessment of a global score that takes into account all symptoms related to this outcome.

In conclusion:

- The frequency, severity and duration of bloating are not appropriate outcome variables to be used alone for the substantiation of claims in the context of reduction of GI discomfort. A SGA of all symptoms combined should be used instead.
- The frequency, severity and duration of bloating are not appropriate outcome variables for the substantiation of claims in the context of maintenance of normal defecation.
- The frequency, severity and duration of bloating are not appropriate outcome variables to be used alone for the substantiation of claims in the context of improved lactose digestion. However, they can

be used as supportive evidence for such health claims.

3.1.1.3.1 QUESTIONNAIRE

See Section 3.1.1.1.1

3.1.1.4 STRAINING

Faecal straining is the contraction of the diaphragm and abdominal wall muscles with a closed glottis. It is a physiological and necessary process during defecation and the straining process has been correlated with stool type. However, change in duration and intensity of straining at stool can be a symptom of various conditions, such as constipation. In this case, prolonged straining may cause hiatus hernia, haemorrhoids, varicose veins in the limbs and deep venous thrombosis (Heaton and Cripps 1993).

To evaluate the appropriateness of straining as OV of reduction of GI discomfort, the literature deriving from database #4 was critically evaluated (Table 1).

The straining forces applied during defecation may be very significant and may let the development of pathological conditions. Straining represents a discomfort for many people (healthy or not) inasmuch it reduces the quality of life, when its duration or severity increase. Other than a pathological conditions, straining can be also a behavioural attitude, given by different situations, i.e. impatience, unfavourable posture while defecating, pelvic floor dyssynergia (anismus), or the sensation of incomplete evacuation (Heaton and Cripps 1993).

Straining often co-exists with one or more of borborygmi, distension, abdominal pain or flatulence. Key symptoms of GI discomfort need to be integrated in a single assessment that it is able to represent an overall effect of the intervention on this outcome. Furthermore, owing the fluctuating nature of GI symptoms, the effect of an intervention should be assessed for extended periods of time (e.g. 4-8 weeks) in order to obtain meaningful results (Irvine et al. 2016, Irvine et al. 2006).

In conclusion, evaluations of the severity and duration of straining cannot be used alone as appropriate outcome variables for the substantiation of claims in the context of reduction of GI discomfort. A SGA of all symptoms combined should be used instead.

3.1.1.4.1 QUESTIONNAIRE

See Section 3.1.1.1.1

3.1.1.5 BORBORYGMI

Borborygmus (plural borborygmi), also known as rumbling or gurgling, is a sound induced by bowel peristalsis, which moves gas through the liquid content of the intestine. Causes of borborygmi may be fasting and incomplete digestion of food leading to an excess of gas in the intestine.

The complete absence of borborygmi may indicate intestinal obstruction, paralytic ileus or other serious pathology.

To evaluate the appropriateness of borborygmi as OV of reduction of GI discomfort, the literature deriving from database #5 was critically evaluated (Table 1).

Borborygmi can be physiological or the result of morbid conditions, such as irritable bowel syndrome (IBS) or celiac disease. In healthy individuals, but mostly in patients with IBS, borborygmi of high severity/frequency induce GI discomfort. The borborygmi are typically associated with other symptoms of GI discomfort such as flatulence, abdominal cramps, bloating and straining, and all of them vary between individuals in frequency and severity. For this reason, key symptoms of GI discomfort need to be integrated in a single assessment that it is able to represent an overall effect of the intervention on this outcome (Spiegel et al. 2010).

In conclusion:

- Borborygmi is not an appropriate outcome variable be used alone for the substantiation of claims in the context of reduction of GI discomfort. A SGA of all symptoms combined should be used instead.
- Borborygmi is not an appropriate outcome variable for the substantiation of claims in the context of maintenance of normal defecation.

3.1.1.5.1 QUESTIONNAIRE

See Section 3.1.1.1.1

3.1.1.6 SENSATION OF COMPLETE/INCOMPLETE EVACUATION

Evacuation is a physiological need which is strictly correlated to the emotional and psychological sphere. In fact, sensations of incomplete evacuation may occur during anxious states or when hygienic conditions are not favourable. Furthermore, constipation or disturbances during defecation may allow the subject to think to an incomplete evacuation. This sensation leads the subject to suffer of pain, intestinal cramps up to an impellent need of evacuate, without any chance of sphincters control.

To evaluate the appropriateness of sensation of complete/incomplete evacuation as OV of reduction of GI discomfort, the literature deriving from database #6 was critically evaluated (Table 1).

The sensation of incomplete evacuation is a subjective symptom associated with GI discomfort. It is difficult to assess because it can vary from patient to patient and from time to time, in severity and duration. It is not a sufficiently validated parameter to be recommended unequivocally as the primary outcome measure for substantiation of health claims related to the reduction of GI discomfort. Furthermore, this symptom interacts with other GI symptoms in complex ways. The feeling of incomplete evacuation is also one of the diagnostic criteria used for the diagnosis of constipation (Stewart et al. 1999).

In conclusion:

- The feeling of incomplete evacuation cannot be used alone as an outcome variable for the scientific substantiation of claims in the context of reduction of GI discomfort. A SGA of all symptoms combined should be used instead.

- The feeling of incomplete evacuation is an appropriate outcome variable for the substantiation of claims in the context of maintenance of normal defecation.

3.1.1.6.1 QUESTIONNAIRE

See Section 3.1.1.1.1

3.1.1.7 ABDOMINAL DISTENSION

Abdominal distension is a visible, measurable, and uncomfortable increase in the abdominal girth. This distension is objectively visible, and it is measurable by several methods, like tape, X-ray,

computed tomography, and abdominal inductance plethysmography. It is usually absent in the morning and progressively appears during the day. Abdominal distension is one of the main features of IBS, although the pathophysiological mechanisms underlying visible distension of the abdomen are not known. It has been hypothesised that abdominal distension may be related to a lower threshold for visceromotor reflexes involved in the regulation of abdominal wall muscle tone, to the increase in intra-abdominal volume due to swallowed air, ingested food and/or fluid, to retained faeces and flatus, and/or to the secretion of digestive juices (Chang et al. 2001, Sullivan 2012).

To evaluate the appropriateness of abdominal distension as OV of reduction of GI discomfort, the literature deriving from database #7 was critically evaluated (Table 1).

Abdominal distension represents a discomfort for many people (healthy or not) inasmuch it reduces the quality of life and it is sometimes associated with pain. During the assessment, it is important distinguish abdominal distension (objective) from bloating (subjective). Abdominal distension is one of the most common and bothersome symptoms in IBS patients. Constipation is characterized by a higher abdominal girth compared to diarrhoea (Agrawal and Whorwell 2008).

In children, abdominal distension is often caused by air swallowing. This discomfort leads children to limit their food intake (Rasquin-Weber et al. 1999).

In conclusion, evaluation of abdominal distention cannot is not appropriate to be used alone as an outcome variable for the substantiation of claims in the context of reducing GI discomfort. A SGA of all symptoms combined should be used instead. Moreover, it is not appropriate to be used alone for the substantiation of such health claims in children.

3.1.1.7.1 QUESTIONNAIRE

See Section 3.1.1.1.1

3.1.1.8 FLATULENCE

Flatulence, also known as farting or passing wind, is the excessive accumulation of air or gas (produced during digestion process) in the intestine that is expelled through the anus, often with sound and/or odour. There are several factors that cause an increase in intensity and occurrence of flatulence,

among which lactose intolerance, malabsorption of certain foods and breakdown of undigested foods due to microbial action. Flatus is predominantly constituted by hydrogen, carbon dioxide, and methane, while the odour is due to other waste trace gases or compounds such as skatole and sulfur-containing substances. Despite these negative aspects related to flatulence, it is a normal biological process and, on average, people have approximately 15 flatus per day (Price et al. 1988, Tomlin et al. 1991).

To evaluate the appropriateness of flatulence as OV of reduction of GI discomfort, the literature deriving from database #8 was critically evaluated (Table 1).

Flatulence represents a discomfort for many people (healthy or not) inasmuch it reduces the quality of life and may become socially disabling when its occurrence or intensity increase. It is commonly a source of embarrassment and can cause distress. Flatulence can have a different aetiology pertaining to physiological or pathological conditions of the gastrointestinal system. However, the mechanisms underlying its physiology and pathophysiology are poorly understood (Manichanh et al. 2014). Flatulence often co-exists with one or more symptoms like borborygmi, distension, abdominal pain or bloating that together lead a decrease in GI comfort. For these reasons, the majority of studies evaluating a reduction of GI discomfort also assessed a reduction of a global score that takes into account all of GI discomfort symptoms during a long period of treatment (e.g. 4-8 weeks)(Irvine et al. 2016, Irvine et al. 2006).

Flatulence is hard to assess. On one hand, people are usually reticent to report on it. On the other hand, individuals may be unaware of flatulence when it occurs because there is either no smell, the amount is tiny, or flatulence is often confused with other symptoms, particularly abdominal bloating. This subjective perception leads to an underestimation of number of gas evacuations (Price et al. 1988). An increase in the severity and occurrence of flatulence may be a symptom of carbohydrate malabsorption, especially lactose. Flatulence appears to be a more reliable indicator of lactose maldigestion than other symptoms. However, there are inter-individual differences in the development of flatulence and cramps in patients with lactose malabsorption, so that the diagnosis of

lactose intolerance cannot rely only on this unspecific symptom (Rao et al. 1994).

In conclusion:

- The intensity and occurrence of flatulence cannot be used alone as outcome variables for the substantiation of claims in the context of reduction of GI discomfort. A SGA of all symptoms combined should be used instead.
- The intensity and occurrence of flatulence are not appropriate parameters for the substantiation of claims in the context of maintenance of normal defecation.
- The intensity and occurrence of flatulence cannot be used alone for the substantiation of claims in the context of improved lactose digestion, but they can be used as supportive evidence.

3.1.1.8.1 QUESTIONNAIRE

See Section 3.1.1.1.1

3.1.1.9 NEED TO DEFECATE/BOWEL URGENCY

Bowel or faecal urgency can be defined as a sudden, irresistible need to have a bowel movement. It is considered an unpleasant sensation as this strong desire to defecate compels people to stop what they are doing and immediately evacuate. Bowel urgency affects about 18% of healthy subjects and 72% of subjects with diarrhoea. Although bowel urgency is most common in patients with diarrhoea-predominant IBS (D-IBS), patients with constipation-predominant IBS and alternating IBS also report faecal urgency (Allen et al. 2004, Basilisco et al. 2007).

To evaluate the appropriateness of need to defecate/bowel urgency as OV of reduction of GI discomfort, the literature deriving from database #9 was critically evaluated (Table 1).

Bowel urgency is a symptom that patients cannot clearly define or describe. To date, the quantification and the characterisation of the urgency sensation perceived by the patient cannot be adequately defined because of the insufficiency of the data. Bowel urgency is not a unidimensional symptom, but rather a multidimensional construct better described by four hierarchically related scales: (i) urgency attributes; (ii) immediacy; (iii) controllability; (iv) psychosocial impact. Due to the difficulty of its evaluation, it should not be considered an appropriate primary endpoint of

605 treatment efficacy in clinical trials. However, bowel urgency represents a symptom clinically
606 meaningful to patients with D-IBS and represents an acceptable co-primary endpoint to assess GI
607 discomfort, if an adequate tool is for its assessment (Spiegel et al. 2010).
608 In conclusion, need to defecate/bowel urgency cannot be used alone as outcome variable for the
609 substantiation of claims in the context of reduction of GI discomfort, because the term "GI
610 discomfort" comprises several symptoms. A SGA of all symptoms combined should be used instead.
611 3.1.1.9.1 QUESTIONNAIRE
612 See Section 3.1.1.1.1
613 3.1.1.10 CONSTIPATION
614 Constipation is a common condition affecting people, especially women, of different ages, such
615 asbabies, children, adults and the elderlies, with a higher prevalence in older adults and during
616 pregnancy According to The North American Society for Pediatric Gastroenterology, Hepatology,
617 and Nutrition, constipation is defined as "a delay or difficulty in defecation, present for two or more
618 weeks, sufficient to cause significant distress to the patient"(North American Society for Pediatric
619 Gastroenterology 2006). According to the Rome III Criteria, constipation takes into account the
620 frequency of defecation, stool consistency, straining and sensation of incomplete evacuation. These
621 symptoms result from a variety of causes, including low dietary fibre intake, emotional or nervous
622 disturbances, structural disorders (such as haemorrhoids, diverticular disease, colon polyps, colon
623 cancer, and inflammatory bowel disease), drug-induced aggravation of constipation, and infections
624 (Alame and Bahna 2012, Arce et al. 2002).
625 To evaluate the appropriateness of constipation as OV of reduction of GI discomfort, the literature
626 deriving from database #10 was critically evaluated (Table 1).
627 Constipation is a disorder of defecation related to bowel habits like stool frequency, consistency, and
628 defecation symptoms. Constipation can lead to bloating and discomfort. This condition reduces the
629 quality of life, both in adults and children. The clinical presentation of constipation includes a broad
630 spectrum of symptoms that are also present in other disorders. Despite constipation is a common

complaint, it is a poorly defined clinical condition (Agachan et al. 1996, Rey et al. 2014). The perception of constipation may include both the objective low stool frequency and subjective alteration of the normal defecation, i.e. faecal straining, incomplete evacuation, abdominal bloating or pain, hard or small stools, or mechanical expulsion of the stools (Arce et al. 2002). Due to the subjective nature of certain symptoms that define constipation, it is important to follow the Rome III Criteria for a correct evaluation of the occurrence and severity of constipation. Diagnostic criteria for constipation must include 2 or more of the following:

a. Straining during at least 25% of defecations;
b. Lumpy or hard stools in at least 25% of defecations;
c. Sensation of incomplete evacuation for at least 25% of defecations;
d. Sensation of anorectal obstruction/blockage for at least 25% of defecations;
e. Manual manoeuvres to facilitate at least 25% of defecations (e.g. digital evacuation, support of the pelvic floor);
f. Fewer than 3 defecations per week.

In conclusion:

- The incidence or severity of constipation are not appropriate to cannot be used alone as outcome variable for the substantiation of claims in the context of reduction of GI discomfort, because the term "GI discomfort" comprises several symptoms. A SGA of all symptoms combined should be used instead. Moreover, these outcomes variables are not appropriate to be used alone for the substantiation of such claims in children.
- The constipation is an appropriate outcome variable for the substantiation of claims in the context of maintenance of normal defecation.

3.1.1.10.1 PATIENT ASSESSMENT OF CONSTIPATION

The Patient Assessment of Constipation (PAC) is a symptom and quality-of-life self-report instrument, composed by two complementary components, the Symptom Questionnaire (PAC-SYM) and the Quality of Life Questionnaire (PAC-QOL), which can be used singularly orin

657 combination.. PAC-SYM is a self- reported questionnaire developed to assess symptom frequency
658 and severity of constipation. This instrument consists of 44 5-point Likert scaled symptom items: 23
659 items assess symptom frequency (none of the time, a little of the time, some of the time, most of the
660 time, and all of the time), and 21 items assess symptom severity (absent, mild, moderate, severe, and
661 very severe). The symptoms included in the questionnaire are: infrequent defecation, pain with
662 defecation, stool size and consistency, straining at stool, sensation of incomplete evacuation, and
663 abdominal pain. The choice of a recall period of two weeks is to limit recall bias and to provide a
664 relatively acute assessment. The PAC-SYM is able to distinguish treatment responders (i.e. subjects
665 who reported in the last visit an improved in severity of constipation, following the treatment) from
666 non-responder (i.e. those patients with no change or some worsening in severity at last visit).
667 Moreover, PAC-SYM is internally consistent, reproducible under stable conditions, valid, and
668 responsive to change, and provides a comprehensive means to assess the effectiveness of a treatment
669 for constipation. The validation of the questionnaire has been conducted on a group of male and
670 female outpatients, aged 18–70 years, with a history of chronic idiopathic constipation in the previous
671 three months. Patients with secondary causes of constipation (for example, endocrine disorders,
672 medication-related) were excluded (Frank et al. 1999).
673 Studies conducted in older people (≥ 65 years) with history of constipation, defined as use of a
674 stimulant or osmotic laxative or enema at least once a week, for the four weeks before the
675 questionnaire, have confirmed the feasibility, acceptability and ease of administration of this
676 questionnaire. A debriefing interview designed to determine whether subjects are able to interpret the
677 meaning of specific terms in the questionnaire correctly should be conducted at baseline. If not, an
678 interviewer-administered PAC-SYM is recommended (Frank et al. 2001).
679 PAC-QOL is a validated Quality of Life Questionnaire developed for use in patients with
680 constipation. Results of validation study demonstrate that the PAC-QOL is internally consistent,
681 reproducible, valid, and responsive to improvements over time (Marquis et al. 2005).
682 In conclusion, the PAC questionnaire is an appropriate method to assess constipation in the general

683 population.

684 3.1.1.11 STOOL CONSISTENCY

685 Analyses of bowel habits and stool characteristics are needed for the diagnosis of diseases which
686 involve changes in normal defecation. Among the various characteristics of faeces, stool consistency
687 is one of the most important. Stools, normally semisolid, could be hard, mucoid, or liquid. Stool
688 consistency may be physiologically modified through diet (e.g. fibre intake) (Halmos et al. 2014,
689 Davies et al. 1986, Deng et al. 2002).

690 To evaluate the appropriateness of stool consistency as OV of reduction of GI discomfort, the
691 literature deriving from database #11 was critically evaluated (Table 1).

692 The analysis of stool consistency is important to identify changes in bowel habits that lead to GI
693 disorders, like constipation or diarrhoea. Hard stools are typical of constipation, with a difficult and
694 painful stool passage through the anus. The stools become hard due to a low water content as a result
695 of low fluid consumption and/or an increased intestinal transit time. A low fibre intake may also lead
696 to hard stools. On the other hand, loose stools are typical of diarrhoea. Soft to watery stools pass out
697 easily and more frequently than normal, and are associated to faecal incontinence. Several studies
698 have shown that changes in stool consistency that lead to a softening of the faeces reduce the risk of
699 constipation, both in adults and children (Bannister et al. 1987). An accurate evaluation of stool
700 consistency required the use a validated method.

701 In conclusion:

702 - Stool consistency is not appropriate to be used alone as outcome variable for the substantiation of
703 claims in the context of reduction of gastrointestinal GI discomfort, because the term "GI discomfort"
704 comprises several symptoms. A SGA of all symptoms combined should be used instead. Changes in
705 stool consistency, however, could be used as evidence in support of the mechanisms by which an
706 intervention may reduce GI discomfort.

707 - Stool consistency is an appropriate outcome variable for the substantiation of health claims in the
708 context of maintenance of normal defecation.

- Stool consistency is an appropriate outcome variable for the substantiation of health claims in the context of contribution to the softening of stools in children.

3.1.1.11.1 BRISTOL STOOL SCALE

The Bristol Stool Scale or Chart is a method to evaluate the stool consistency. This seven-point scale was validated in healthy control subjects and in patients with GI disorders. Its efficacy and reliability in discriminating between healthy individuals and individuals with pathological conditions affecting stool consistency has been demonstrated clinically and for research purposes. It recognises seven types of stools:

Type 1: Separate hard lumps, like nuts; Hard to pass;

Type 2: Sausage shape but lumpy;

Type 3: Like a sausage but with cracks on the surface;

Type 4: Like a sausage or snake, smooth and soft;

Type 5: Soft blobs with clear cut edges; Passed easily;

Type 6: Fluffy pieces with ragged edges; A mushy stool;

Type 7: Watery with no solid pieces. Entirely liquid.

The Bristol Stool Scale incorporates images illustrating faecal samples, along with precise descriptions of the shape and consistency of stools, using easily recognizable examples (Martinez and de Azevedo 2012, Pares et al. 2009).

A modified Bristol Stool Scale was created and validated for use in children (Lane et al. 2011).

In conclusion, the Bristol Stool Scale appears to be a reliable and appropriate technique for measuring stool consistency, both in adults and children.

3.1.1.12. DIARRHOEA

Diarrhoea is more a symptom than a disease and can be present in many different conditions, like IBS, celiac disease, Crohn's disease, GI infections and lactose intolerance. Diarrhoea is characterised by loose or watery stools and is most common in children. Diarrhoea may have subjective meanings

735 and most patients consider loose stools as the key characteristic of diarrhoea. However, this symptom
736 is characterized by many other factors. It is usually defined as three or more loose or watery stools in
737 a 24-hour period and can be classified as acute (lasting < 2 weeks) or persistent (lasting 2 weeks or
738 more). For infants, the definition of diarrhoea is different than for adults, because loose stool pass
739 more frequently in normal conditions, especially in infants who are breastfed. For this reason, the
740 diagnosis of diarrhoea in infants is made by the mother on the basis of what is abnormal for her child
741 (Lee et al. 2012). However, the physician's diagnosis is necessary for research purposes. Collateral
742 effects of diarrhoea are dehydration and dysentery.
743 To evaluate the appropriateness of diarrhoea as OV of reduction of GI discomfort, the literature
744 deriving from database #12 was critically evaluated (Table 1).
745 Diarrhoea is considered a defecatory symptom and it is used to assess changes in bowel habits.
746 Beneficial changes in bowel habits should not lead to diarrhoea.
747 There are many factors that need to be taken into consideration in order to define diarrhoea, like the
748 frequency, duration and severity of diarrhoea episodes. There are many causes of diarrhoea, which
749 may be infectious or not. A number of non-infectious medical conditions may cause diarrhoea, for
750 example lactose maldigestion, celiac disease, IBS, inflammation of the bowel, use of antibiotics or
751 cancer. Regarding lactose maldigestion, symptoms like diarrhoea do not show a significant
752 relationship with breath hydrogen excretion, which is considered the gold standard method for the
753 assessment of lactose maldigestion (Hammer HF et al 2012) (Rao DR et al 2014).
754 In conclusion:
755 - The frequency, severity and duration of diarrhoea are not appropriate outcome variables to be used
756 alone for the substantiation of claims in the context of reduction of GI discomfort, because the term
757 “GI discomfort” comprises several symptoms. A SGA of all symptoms combined should be used
758 instead. Moreover, the outcome variables are not appropriate to be used alone for the substantiation
759 of such claims in children.
760 - The frequency, severity and duration of diarrhoea are appropriate outcome variables for the

761 substantiation of claims in the context of maintenance of normal defecation.

762 - The frequency, severity and duration of diarrhoea are not appropriate outcome variables to be used
763 alone for the substantiation of claims in the context of improved lactose digestion. However, they can
764 be used as supportive evidence for such health claims.

765 3.1.1.12.1 QUESTIONNAIRE

766 See Section 3.1.1.1.1

767 3.1.1.13 STOOL FREQUENCY

768 Stool frequency, also known as frequency of bowel movements or frequency of defecation, is the
769 frequency whereby the stool passes through the anus, without manual manoeuvres or rescue laxatives.
770 A physiological bowel frequency varies from two to three times per day to once every three days
771 (Heaton et al. 1992), while diarrhoea or constipation occurs when defecation is, respectively, more
772 or less frequent than that (Longstreth et al. 2006). In most cases, changes in stool frequency are not
773 as sign of disease, but rather an indicator of a change in dietary habits, routine, stress levels or even
774 physical exercise. They may also be associated with the use of stimulants, like nicotine or caffeine,
775 especially if there is excessive use within a short period of time.
776 To evaluate the appropriateness of stool frequency as OV of reduction of GI discomfort, the literature
777 deriving from database #9 was critically evaluated (Table 1).
778 The normal length of time between bowel movements ranges widely from person to person. Stool
779 frequencies outside the physiological ranges can occur. In children, the most important factor that
780 affects the frequency of defecation is the children age. The most frequent defecation occurs in the
781 first month of life and decreases with increasing age (Weaver and Steiner 1984). The analysis of the
782 frequency of bowel movements is important to identify changes in bowel habits that can lead to GI
783 disorders, like constipation (stool frequency is often used to define constipation, but as the sole
784 criterion it may not be sufficiently comprehensive) or diarrhoea. If more than three days pass without
785 having a bowel movement, the stool becomes harder and more difficult to pass, which may cause
786 pain and discomfort. In some studies involving children, a correlation is shown between low

frequency of bowel movements (less than once a day) and presence of hard stools (Weaver and Steiner 1984). However, there are cases of frequent bowel movements that cannot fit into the classical presentation of diarrhoea and an increase in stool frequency may not lead to changes in the consistency or colour of the faeces.

In conclusion:

- Stool frequency is not an appropriate outcome variable for the substantiation of claims in the context of GI discomfort, because the term "GI discomfort" comprises several symptoms. A SGA of all symptoms combined should be used instead. Changes in stool frequency, however, could be used as evidence in support of the mechanisms by which an intervention may reduce GI discomfort.
- Stool frequency is an appropriate outcome variable for the substantiation of claims in the context of maintenance of normal defecation.
- Stool frequency is not an appropriate outcome variable for the substantiation of claims in the context of contributing to softening of stools in children, but it can be used as supportive of a mechanism through which the food/constituent could exert the claimed effect.

3.1.1.13.1 QUESTIONNAIRE

See Section 3.1.1.1.1

3.1.1.14 QUALITY OF LIFE

Quality of life is a generic and broad term, the definition of which depends on a variety of factors, including the support from friends and relatives, the ability to work and be interested in its own occupations, as well as health and disabilities. Health-related quality of life is a concept encompassing illness experience, functional status and the perceptions of the subject related to a medical condition. Social, cultural, psychological and disease-related factors have an effect on it (Felce and Perry 1995). GI discomfort can negatively impact the quality of life up to compromise it in case of severe symptoms.

To evaluate the appropriateness of quality of life as OV of reduction of GI discomfort, the literature deriving from database #13 was critically evaluated (Table 1).

The measurement of health-related quality of life allows a composite evaluation of the patient's condition, which results from biological (objective) and psychological (subjective) factors. Investigators can use the assessment of health-related quality of life to compare the data across subject cohorts, but also to evaluate the response to a treatment in intervention studies (Wong and Drossman 2010). In addition, there can be some discrepancies between the patient's and the physician's perception in relation to the success of a treatment that aims at improving the symptoms of a disease, rather than at curing the disease. In these cases, it is suitable to measure the success of the treatment in terms of the improvement of the quality of life of the patient.

In conclusion:

- The quality of life cannot be used alone as an outcome variable to substantiate of health claims referring to the reduction of GI discomfort. However, it can be used a supportive evidence.
- The quality of life cannot be used alone as an outcome variable to substantiate health claims referring to the maintenance of normal defecation. However, it can be used a supportive evidence.

3.1.1.14.1 FUNCTIONAL DIGESTIVE DISORDERS QUALITY OF LIFE QUESTIONNAIRE

Functional digestive disorders quality of life questionnaire (FDDQOL), developed by Chassany et al. with the aim of providing a measure of the quality of life for patients with functional dyspepsia and IBS, is one of the first functional, digestive disease-specific Quality of Life (QOL) tools (Chassany et al. 1999). The original 74 items have been subsequently reduced to 43 on its current version. The FDDQOL has 8 domains: daily activities (8 items), anxiety (5 items), diet (6 items), sleep (3 items), discomfort (9 items), coping with disease (6 items), control of disease (3 items) and stress (3 items). Referring to their condition over the past fortnight, the subjects assign individual scores to each item using a 6-point Likert scale as response format. The score for each scale is then obtained by the sum of the scores for each item and transformed into a scale from 0 to 100 corresponding to the worst and the best possible health state measured by the questionnaire, respectively. Finally, a global score (ranged from 0 to 100) is computed from the scale scores. The questionnaire has shown good

reliability. Compared to a generic QOL tool, FDDQOL has demonstrated concurrent validity. Even if the psychometric quality is good, a consensus panel found it of insufficient methodological quality and practical utility (Wong and Drossman 2010).

In conclusion, FDDQOL can be considered an appropriate method to assess the quality of life in individuals with GI symptoms, provided that its limitations are taken into consideration.

3.1.1.14.2 IRRITABLE BOWEL SYNDROME-36

Irritable bowel syndrome-36 (IBS-36) represents an IBS-specific health related QOL questionnaire designed to be self-administered by subjects suffering from this syndrome. The first version of the questionnaire had 70 items divided into 8 domains: daily activities, emotional impact, family relations, food, sleep and fatigue, social impact, sexual relations and symptoms. Subsequently, trough statistical and consensus methodologies, the number of items was reduced to 36. The score is done on a 7-point Likert scale ranging from 0 (symptom never occurred) to 6 (symptom always occurred) corresponding to best and worst quality of life respectively, with a maximum final score of 216. IBS-36 is a retrospective tool, with a recall period of the preceding two months. The questionnaire presents high level of internal consistency and test-retest reliability. IBS-36 allows an evaluation of specific symptoms and areas of the disease that have an impact on the subject's health-related QOL (Groll et al. 2002). Unlike generic instruments, the disease-specific IBS-36 is not helpful outside the target population of IBS patients for which it was developed. On the other hand, generic heath related QOL questionnaires are not specifically addressed to measure gastrointestinal symptoms. Thus, they may be insensitive to changes associated to IBS and are not appropriate to fully capture health related QOL as outcome variable in patients with IBS before and after an intervention (Wong and Drossman 2010).

In conclusion, IBS-36 questionnaire is an appropriate method to assess health-related quality of life in patients with IBS.

3.1.1.14.3 SHORT FORM-36

SF-36 is a generic and short health-related QOL questionnaire which comprises 36 items evaluating

nine domains: physical and social functioning (10 and 2 items, respectively), role limitation by physical and emotional problems (4 and 3 items, respectively), mental health (5 items), energy and vitality (4 items), bodily pain (2 items), general perception of health (5 items), and changes in health over the past year. The latter domain is an unscaled single item. The answering options can be dichotomic or relate to three, five, or six-point Likert scales. For each variable, item scores are coded, summed and transformed into a scale from 0 to 100 corresponding to the worst and the best possible health state measured by the questionnaire, respectively (Jenkinson et al. 1993). SF-36 is a retrospective tool, with a recall period of the preceding four weeks. It is found acceptable by the patients and shows high levels of internal validity and good test-retest properties. The response rate for SF-36 has been found to be different for different age groups. Lower response rates have been reported among people aged 75 years and over with poor physical/mental health scores, because of inability to self-complete the questionnaire. The main reasons are associated to visual impairment or writing difficulties. Furthermore, some questions related to work or physical activity, not specifically developed for these subjects, can be easily missed (Hayes et al. 1995). These aspects should be considered when using this tool that can be potentially useful for measuring health status in medical research. Various forms of SF-36, some of which have not been validated, are currently available.

In conclusion, validated versions of SF-36 are appropriate method to assess health related QOL. However, since it is not specifically addressed to measure gastrointestinal symptoms, it may be insensitive to changes associated to IBS and it is not appropriate to fully capture health related QOL as outcome variable in patients with IBS before and after an intervention.

3.1.1.14.4 RAND 36-ITEM HEALTH SURVEY

The RAND 36-item health survey is a generic health-related QOL instrument widely used in the world. The instrument has 8 health domains with a total of 35-item scales: physical and social functioning (10 and 2 items, respectively), role limitations caused by physical health and emotional problems (4 and 3 items, respectively), emotional well-being (5 items), energy/fatigue (4 items), pain (2 items) and general health perception (5 items). Physical and mental health summary scores are

891 derived from these scales. The remaining item assesses change in the perception of health in the last
892 12 months (Hays and Morales 2001). The RAND survey includes the same items as the SF-36 but
893 uses a two-step process for scoring. Equivalent results are obtained for 6 of the 8 subscales, with
894 different scoring for pain and general health perception scales. Rand questionnaire requires only 7-
895 10 minutes to be filled. It can be filled by subjects, or administered by the investigator during a
896 telephone personal interview. Questionnaires administered by e-mail are cheaper, but the response
897 rate and completeness are lower than by phone (Hays and Morales 2001).
898 In conclusion, the Rand-36 item heath survey is an appropriate method to measure health related
899 QOL. However, since it is not specifically addressed to measure gastrointestinal symptoms, it may
900 be insensitive to changes associated to IBS and it is not appropriate to fully capture health related
901 QOL as outcome variable in patients with IBS before and after an intervention.

902 3.1.1.15 COMPOSITION OF THE GUT MICROBIOTA/ BIFIDOBACTERIAL
903 POPULATION

904 The human gut is a natural reservoir for numerous species of microorganisms and contains ~1×10^{12}
905 bacterial cells/g of colonic content. More than 500 bacterial species populate the gut of healthy
906 individuals, with predominance of obligate anaerobes, located mainly in the colon. The dominant
907 phyla are Actinobacteria, Firmicutes, Proteobacteria and Bacteroidetes. The mutualistic relationship
908 between symbionts and commensals, and the diversity and stability of microbiota, are important for
909 the maintenance of health and wellbeing; alterations in this balance or diversity leads to dysbiosis,
910 and ultimately to clinical disease expression (Malinen et al. 2005). Humans are colonized at birth and
911 development of microbiota is influenced by many factors such as type of birth, gestational age, use
912 of antibiotic and feeding. The microbiota evolves during different stages of life (Gareau et al. 2010).
913 To evaluate the appropriateness of composition of the gut microbiota/bifidobacterial population as
914 OV of reduction of GI discomfort, the literature deriving from database #14 was critically evaluated
915 (Table 1).
916 The gut microbiota ensures normal bowel physiological functions, works as a barrier against

pathogens and stimulates the host immune function by releasing different metabolites and chemicals (e.g. butyrate, which is essential for the integrity of the colonic epithelium). Some studies suggest that the gut microbiota could play a role in the pathogenesis of IBS (Collins 2014). Scientific evidence demonstrates that the diversity, stability and metabolic activity of the gut microbiota are compromised in subjects with some diseases (e.g. with inflammatory bowel disease, irritable bowel syndrome, obesity, diarrhoea, necrotizing enterocolitis) compared to healthy individuals, but little is yet known about the health relevance of individual microbial species or strains. The gut microbiota is also critical for the maturation of the host's mucosal immune system during early life and this function continues throughout life. Moreover, developmental aspects of the adaptive immune system are influenced by bacterial colonization of the gut (Gareau et al. 2010).

However, despite the emerging evidence linking the composition of the gut microbiota to GI disease and immune function, changes in the composition of the gut microbiota do not describe a specific function of the body.

In conclusion, the composition of the gut microbiota is not an appropriate outcome variable to be used:

- For the substantiation of health claims in the context of reduction of GI discomfort.
- For the substantiation of health claims in the context of maintenance of normal defecation.
- For the substantiation of health claims in the context of initiation of appropriate immune responses, including the defence against pathogens.

However, changes in the composition of the gut microbiota could be used in support of the mechanisms by which the food/food component may exert these claimed effects.

3.1.1.15.1 16S rRNA MICROBIAL PROFILING

16S rRNA microbial profiling is a key tool for studies of microbial communities. The 16S rRNA gene, contained in the nuclear DNA, codifies for the ribosomal RNA which is part of the small subunit of the ribosomes. It represents a molecular marker widely used in bacterial taxonomy because of its conservations despite the evolution of the species. This analysis exploits the recent applications of

metagenomics in the field of microbial ecology. Briefly, this method consists of the extraction of bacterial DNA from a biological sample (faeces or intestinal biopsy) and the subsequent amplification of the 16S rRNA gene with an appropriate primer pair. The analysis is completed by sequencing the 16S rRNA gene PCR products corresponding to the microorganisms present in the microbiota. Finally, by the use of bioinformatics tools, it is possible to recognize the exact composition of gut microbiota, identifying also the microorganisms that are not cultivable, and observe changes in the gut microbiota composition (at the level of genera). This validated method is highly reproducible and has a high throughput. However, 16S rRNA microbial profiling of the human gut microbiota is strongly influenced by sample processing and PCR primer choice. Therefore, appropriate primer selection as well as DNA extraction protocols are essential to enable trustworthy representation of the organisms present in an environment, such as the human gut ecosystem (Milani et al. 2013).

In conclusion, 16SrRNA microbial profiling is an appropriate method to assess the composition of the gut microbiota.

3.1.1.15.2 BIFIDOBACTERIA ITS profiling

This method exploits the great deal of sequence and length variation of ITS (Internal Transcribed Spacer) regions, which is useful for differentiating species of prokaryotes.

The sequences of the spacer region are comprised between the 16S rRNA and the 23S rRNA genes within the rRNA locus. The method consists of the extraction of bacterial DNA from a biological sample and subsequent amplification of the ITS regions with appropriate primer specific for Bifidobacteria. The analysis is completed by sequencing the amplified regions. With the use of bioinformatics tools, it is possible recognize the exact composition of Bifidobacteria species. ITS sequence analysis is a useful technique for identifying Bifidobacteria at the species level (Milani et al. 2014).

In conclusion, Bifidobacteria ITS profiling is an appropriate method to assess bifidobacterial population of the gut.

3.1.2. REDUCTION OF EXCESSIVE INTESTINAL GAS ACCUMULATION

3.1.2.1 INTESTINAL GAS VOLUME

The most important gases in the human gut are nitrogen (N_2), oxygen (O_2), hydrogen (H_2), carbon dioxide (CO_2), and methane (CH_4). Instead, hydrogen sulphide (H_2S), methanethiol (CH_3SH), and dimethylsulphide (CH_3SCH_3) are present in trace (1%) and they are responsible for the characteristic unpleasant odour of intestinal gas (Suarez and Levitt 2000). Gas is introduced into the gastrointestinal tract in several ways. In particular, there are four main mechanisms that deliver gases to the intestinal lumen: 1) air swallowing (O_2 and N_2); 2) interaction of bicarbonate and acid (CO_2); 3) diffusion from the blood (CO_2, N_2, and O_2); and 4) bacterial metabolism (CO_2, H_2, CH_4, and sulphur-containing gases). These gases are then eliminated from the gut through oesophagus (belching) or anus (flatulence), or diffusion into the blood. The set of these processes determines the volume and mean composition of the entire gastrointestinal gas.

To evaluate the appropriateness of intestinal gas volume as OV of reduction of GI discomfort, the literature deriving from database #15 was critically evaluated (Table 1).

In the fasting state, the healthy GI tract contains about 100 mL of gas (mean of 100 mL, maximum of 200 mL). The volume of gas increases by about 65% during the postprandial period, primarily in the pelvic colon, with no significant gas accumulation in other gut compartments (Pritchard et al. 2014). Several factors, including gastrointestinal and non-gastrointestinal diseases, dietary habits, and side effects of various drugs, may lead to an accumulation of intestinal gas increasing its volume. Therefore, from such observations it can suppose a strong correlation between gas accumulation and gas volume. In fact, a reduction of one leads to the decrease of the other.

Furthermore, the excessive volume of intestinal gas can be the cause of bloating and distension, but this link has not been yet ascertained. In fact, the only available results suggest that increased gas volume may not be the main mechanism of bloating, but rather impaired gas transit or distribution are more often the cause of problem (Suarez and Levitt 2000).

In conclusion, the measurement of the intestinal gas volume is an appropriate outcome variable for the substantiation of health claims regarding the reduction of excessive intestinal gas accumulation.

995 Moreover, this outcome variable is appropriate for the substantiation of health claims in the context
996 of the reduction of GI discomfort in children.

997 3.1.2.1.1 MAGNETIC RESONANCE IMAGING

998 Magnetic Resonance Imaging (MRI) is a highly sophisticated and most costly technique, now
999 extensively used in body composition research and it is able to measure intestinal gas volume
1000 (Pritchard et al. 2014). In fact, some data suggest the potential use of MRI to estimate the amount of
1001 gas in the gut (proving an excellent accuracy in evaluate intestinal gas volume), which represents a
1002 crucial issue in patients with IBS and other GI disorders with abnormal gas dynamics (Lam et al.
1003 2017). Additionally, MRI may facilitate assessment of the effect of drugs on gas production and
1004 transit within the gut. MRI process requires a magnet, usually a superconducting one, a magnetic field
1005 gradient system for signal localization and a radio frequency system, which is used for signal
1006 generation and processing. The array data provided by MRI, as well as other imaging techniques,
1007 shows the spatial distribution of physical quantities and gas appears as signal void within the bowel.
1008 There are multiple methods to determine gas volume with MRI, including extrapolation of single-
1009 slice or multiple slice acquisitions from both selected regions of the body or the whole-body
1010 measurements. Since this is a time-consuming technique, single-slice imaging is often chosen, in spite
1011 of being less accurate. However, whole-body scans necessarily need to be acquired as a series of
1012 stacks and then integrated. Currently, the accuracy of MRI can be limited by the size pixels (2 mm x
1013 2 mm) employed in bowel scan, as well as by the image distortion deriving from the use of images
1014 obtained with multi-slice technique. A commonly used approach is manual or semi-automated
1015 analysis of time intensive T1-weighted images. Measurements are operator-dependent in case of
1016 manual input. The majority of the automated validated procedures have dealt with the assessment of
1017 adult subjects. Due to reduced compliance of children with the MRI technique, which requires small
1018 movements and sometimes breath holding, measurements in these subjects are rather complicated.
1019 Additional reasons for reduced accuracy in children are their small body size. Moreover, this
1020 procedure is safe because it does not expose subjects to ionizing radiations. However, MRI might not

be a suitable method for routine fieldwork in large-scale studies and its limitations are mainly due to costs.

In summary, MRI is an appropriate method to assess intestinal gas volume.

3.1.2.2 BREATH HYDROGEN CONCENTRATION

The colonic microbiota contains more than bacterial species and plays an important role in human digestive physiology. Most of these microorganisms are saccharolytic and the products of fermentation of dietary carbohydrates are mainly short-chain fatty acids (acetic, propionic and butyric acid) and gases (CO_2, CH_4 and H_2) (Perman et al. 1984). In particular, hydrogen gas (H_2) is produced in the lumen of the gastrointestinal tract. This gas either passes as flatus, or diffuses into the body and is exhaled. In fact, some of the hydrogen produced by the bacteria, whether in the small intestine or the colon, is absorbed into the blood flowing through the wall of the small intestine and colon. The hydrogen-containing blood travels to the lungs where the hydrogen is released and exhaled in the breath where it can be measured.

To evaluate the appropriateness of breath hydrogen concentration as OV of reduction of GI discomfort, the literature deriving from database #16 was critically evaluated (Table 1).

The intestinal gas is composed by five major components: O_2, N_2, H_2, CO_2 and CH_4. Since the latter three are not found in inhaled air, they must be produced in the gut. Several factors lead to an increase of intestinal gases among which:

- Swallowed air;
- SIBO (Small Intestinal Bacterial Overgrowth);
- Breakdown of undigested foods;
- Maldigestion or malabsorption of sugars and polysaccharides (e.g., lactose intolerance).

The bowel contains an enormous number of bacteria that are predominantly anaerobes and produce a large quantity of gases, mainly hydrogen (Rana and Malik 2014). The hydrogen generated in the intestine is absorbed into the portal circulation and excreted in breath. There is strong evidence that the exhaled hydrogen indicates the quantity and the metabolic activity of anaerobic bacteria in the

intestine. Although gas accumulation is one of the major symptoms of GI discomfort (it causes pain and bloating), both in adults and in children, the use of breath H2 tests to evaluate intestinal gas accumulation, has limited specificity and sensitivity.

Instead, for the detection of carbohydrate malabsorption, the measurements of breath hydrogen (H_2) are widely used in clinical medicine. In particular, for the diagnosis of fructose or lactose malabsorption, lactose maldigestion (reduced enzymatic capacity to digest lactose) as well as for the detection of small intestinal bacterial overgrowth syndrome, the hydrogen breath level is widely measured because is considered to be the most reliable outcome, provided the suitable substrates (e.g. lactose for evaluate lactose maldigestion) are used.

However, about 15%-30% people are considered non-H_2 producers because of the presence of Methanobrevibacter smithii in their gut microbiota (Mathur et al. 2013). Since it metabolizes four atoms of hydrogen to form one molecule of methane, an increase in H_2 levels in breath is not observed. In these patients, it is necessary to carry out a lactulose test. If a lactulose load still does not produce an increase in H2 levels, the subject is very likely to be a non-H_2 producer.

In conclusion:

- The levels of breath hydrogen is not an appropriate outcome variable to be used alone for the substantiation of health claims related to the reduction of excessive intestinal gas accumulation (generally leads to a reduction in GI discomfort), but it can be used as supportive outcome.
- The levels of breath hydrogen is an appropriate parameter for the substantiation of health claims in the context of improving of lactose digestion, provided that are performed by appropriate techniques for a correct evaluation.
- The levels of breath hydrogen is not an appropriate outcome variable to be used alone for the substantiation of health claims related to the reduction of GI discomfort in children, but it can be used as supportive outcome.

3.1.2.2.1 BREATH HYDROGEN TEST

Breath hydrogen test is a method that uses the measurement of H_2 in the breath to diagnose several

conditions that cause gastrointestinal symptoms. It is based on the physiological fact that healthy humans when fasting and at rest do not exhale H_2. Breath hydrogen test is used in the diagnosis of carbohydrates malabsorption, SIBO and to assess the orocecal transit time (OCTT) (Rana and Malik 2014). The breath test is preceded by a fasting period of 12 hours; then, the test starts with the blowing into a balloon, which allows the quantification of the basal H_2. The patient then ingests a small amount of the test sugar (lactose, sucrose, sorbitol, fructose, lactulose, etc. depending on the purpose of the test). While glucose breath hydrogen test is more specific SIBO diagnosis, lactose and fructose breath hydrogen tests are used for lactose and fructose maldigestion diagnosis, respectively. Lactulose breath hydrogen test is also used widely to measure the OCTT for GI motility. Every 15 minutes, for up to five hours, H_2 is measured and, in general, an increase in H_2 concentrations of more than 20 ppm above the basal value is considered to be a positive test result. In certain people, it is possible to obtain false-negative results, due to the inability of colonic flora to produce H_2 (non-H_2-producer), or after a recent use of antibiotics or due to a longer orocecal transit time. A more precise diagnosis of non-H_2-production may be done by performing a lactulose test and, if a slow transit time is suspected, it is recommended to do additional readings and extend the test.

False-positive breath tests are less frequent and are mainly due to small bowel bacterial overgrowth or abnormal oral microbiota for this is recommended brush the teeth prior the test.

Although some problems, adopting precautions and following precise guidelines for the interpretation of the results may help to improve the quality and reliability of the test. The lactulose hydrogen breath test is non-invasive, low cost and it can be applied both in adults and children (except for sorbitol and xylitol tests). For the diagnosis of fructose or lactose malabsorption and SIBO, hydrogen breath test is considered the gold standard. Moreover, the lactulose hydrogen breath test allows accurate measurement of orocecal transit time if a hydrogen threshold increment of 5 ppm is chosen.

In conclusion:

- The breath hydrogen test is the most appropriate method for evaluating level of hydrogen in breath, in both adults and children.

- The breath hydrogen test is an appropriate method for evaluating intestinal transit time.

3.1.3 MAINTENANCE OF NORMAL DEFECATION

3.1.3.1 STOOL FREQUENCY

See Section 3.1.1.13

3.1.3.1.1 QUESTIONNAIRE

See Section 3.1.1.13.1

3.1.3.1.2 DIARY

Diaries are a method developed for minimize recall bias and capture experiences close to the time of occurrence, for evaluation of several endpoints. The main advantage of these prospective tools is that they are not affected by the memory, differently from retrospective methods (e.g. recall), and this can be particularly important for the elderlies (Lackner et al. 2014). However, a major problem of diaries is poor adherence, in that patients who failed to complete them, or complete them retrospectively. For these reasons, there are concerns about compliance of paper diaries (McColl 2004). However, electronic device with reminder alarm can improve adherence.

In conclusion, electronic diary and not paper diary, can be an appropriate method to assess stool frequency, as well as other GI symptoms.

3.1.3.2 STOOL CONSISTENCY

See Section 3.1.1.11

3.1.3.2.1 BRISTOL STOOL SCALE

See Section 3.1.1.11.1

3.1.3.3 STOOL WEIGHT/VOLUME/SIZE

The term “faeces” means the remaining material after food is digested along with water, bacteria and other substances secreted into the gastrointestinal tract. There are many characteristics to describe faeces among which stool weight (Myo et al. 1994). It depends mainly on the presence of water, bacteria and fibre in the faeces. About 75% of faecal weight is made up of unabsorbed water (contributing to wet faecal weight). The remaining 25% is composed of solid matter that contains

principally bacteria (responsible for half of the dry faecal weight) as well as undigested fibre and solidified components of digestive juices, fat, inorganic matter and protein. Indicatively, people who consume fibre-rich diets excrete up to 400 grams of stools daily.

To evaluate the appropriateness of stool/weight/size as OV of maintenance of normal defecation, the literature deriving from database #11 was critically evaluated (Table 1).

In a healthy subject, diet quality and quantity are important determinant of stool weight, as, for example, a diet rich in fibre can provide an increase in increase in the daily stool weight, while it can be reduced by a diet rich in fat (Cummings 2001). Other factors able to affect stool weight are sex, ethnicity and body weight (Rose et al. 2015). Furthermore, stool weight varies markedly among different populations, being relatively low in developed countries and also depends on race, ethnicity and dietary habits. Besides all this, it is not known if the stool weight can be a valid parameter to evaluate the severity of various discomforts associated with bowel movements and there are limited data on stool weight among healthy subjects.

In conclusion:

- Stool weight/volume/size is not an appropriate outcome variable to be used alone for the substantiation of claims in the context of maintenance of normal defecation. However, it can be used as supportive evidence to substantiate such health claims.
- Stool weight/volume/size is not an appropriate outcome variable to be used for the substantiation of claim in the context of softening of stool in children.

3.1.3.3.1 DIRECT ASSESSMENT BY THE INVESTIGATORS

The best method to evaluate stool weight is the weight performed by researches using a laboratory scale. Hygiene pads are usually used for collection of hard stool, while stool collectors are used in case of watery or loose stools. The faecal material is then transferred in pre-weighed buckets and weighed on a laboratory balance. The balance need to be calibrated and suitable for use (analytical balance). It is accepted a minimal leakage of faeces that is the soiling on the toilet paper.

The stool can be stored at 4°C for 1 day before the weighing.

In conclusion, the direct assessment by the investigators, represents an appropriate method to evaluate stool weight/volume/size.

3.1.3.4 INTESTINAL TRANSIT TIME

Intestinal motility is a critical process underlying the major functions of the bowel such as storage, absorption, propulsion and defecation. Disorders of colonic motility typically occur with constipation or diarrhoea. Intestinal transit time is useful in evaluating intestinal motility since it represents the length of time taken by food to move through the digestive tract (Spiller 1994). Once food is chewed and swallowed, it moves to the stomach, where it is mixed with acid and digestive enzymes. Subsequently, the food is squeezed through the small intestine, where nutrients are absorbed. The food then moves to the colon: here undigested and unabsorbed food from the small intestine combine with bacteria for the colic fermentation and digestion. After this last passage, together with other waste products, stools are formed, and they are ready to be expelled through the anus.

To evaluate the appropriateness of intestinal transit time as OV of maintenance of normal defecation, the literature deriving from database #17 was critically evaluated (Table 1).

Disturbances in motility and transit are common in functional gastrointestinal (GI) disorders such as irritable bowel syndrome, functional dyspepsia, gastroparesis, bloating or chronic idiopathic constipation (Kusano et al. 2014). One of the main drawbacks of the diagnosis is the difficulty in understanding of which GI region is affected because of the symptoms, which are in common with several other discomforts. However, the assessment of transit through the GI tract provides useful information regarding gut physiology and pathophysiology and allows to evaluate the severity of the problem and help in formulating the diagnosis and the prognosis. The ideal intestinal transit time is from 12 to 24 hours. When these times are exceeded, risk of diverticulosis and candidiasis as well as inflammation and cancer are increased. Furthermore, toxins and wastes may be driven back into the bloodstream, causing, headaches, gas, bloating, acne, allergies, muscle and joint pain. On the contrary, a GI transit time shorter than 10 hours may counteract the normal absorption of nutrients

from food. Thus, besides nutritional deficiencies, electrolyte imbalances, anaemia and osteoporosis may occur. However, bowel transit time is also influenced by the type of food eaten, hydration, the amount of dietary fibre, and exercise. For example, people who eat high amounts of fruits, vegetables and whole grains tend to have a shorter transit time than those who eat mostly sugars and starches. Certain medications (e.g. cold medicines, iron, or medicine used to control blood pressure and pain) and several diseases (e.g. hypothyroidism, diabetes, or Hirschsprung's disease) can also affect transit time contributing to constipation or to loose stools (Tack and Janssen 2010). Furthermore, intestinal transit time varies markedly among different populations. It depends on race, ethnicity and dietary habits. The methods for the measurement and standardized protocols for one population may not be applicable to another population. Intestinal transit time should be standardized and validated for the individual population. As different people have different transit times depending on several factors, intestinal transit time testing is not recommended to evaluate bowel habits.

In conclusion, intestinal transit time is not an appropriate outcome variable to be used alone for the substantiation of claims in the context of maintenance of normal defecation, but it can be used as supportive of the mechanisms by which the food/food component may exert the claimed effect.

3.1.3.4.1 ROM TECHNIQUE

The use of radio-opaque markers (ROM), followed by abdominal X-rays, is a method used to measure total and segmented CTT (colonic transit time) and WGTT (whole-gut transit time) (Ghoshal et al. 2007). This is a quantitative method where radio-opaque pellets are used as probes. This technique has the advantage that the probe can be detected by taking serial radiographs of the abdomen/stools and these pellets are easy to take (swallow with a drink). Following the disappearance of the markers from the gut or the appearance of the markers in the stool by radiographs is required to evaluate transit time. Radio-opaque markers have a well-established role in distinguishing between patients with normal and those with slow intestinal transit, but in the latter group their accuracy in defining the region of delay has not been established, especially if no frequent radiographs are performed. In contrast, daily radiographs involve a high dose of radioactivity (van der Sijp et al. 1993). Intrinsic

drawbacks of the ROM test include radiation exposure (especially for children and patients in child-bearing age), inability to assess regional gut transit, and lack of standardized protocols for the test/interpretation. Also, although some protocols require multiple visits which affect compliance, the ROM technique is commonly used for measuring colonic transit and is often used as gold standard, even if there is no universally accepted or standardized technique for assessing CTT and WGTT. However, the measure of transit time by ROM, can be performed with reasonable accuracy by administration of 10–12 radiopaque markers daily for 6 days, followed by made a radiography on day 7. Following this procedure, this method can be recommended for use in clinical practice and in research.

In conclusion, ROM technique may be an appropriate method for assessing intestinal transit time.

3.1.3.4.2 SST WITH COLOURED PLASTIC PELLETS

A method for the assessment of intestinal transit time is the single stool transit (SST) with the use of coloured plastic pellets (Stevens et al. 1987). These markers must be in different colours and are 3-4 mm in length and 1 mm in diameter and have a specific gravity of about 1.3. These pellets (about 100 markers/day, 20 for each colour) are administered for 3 days, though 6 days are better. The pellets are recovered from the stool by visual inspection and sample number one is the first stool passed 3 h after the last dosing., although it is non-invasive, this method has several limitations, including the inability to monitor pellet transit through the intestinal tract and the possibility of not recovering all the pellets due to errors in sifting the faeces.

In conclusion, SST with coloured plastic pellets is not an appropriate method to assess intestinal transit time.

3.1.3.4.3 BREATH HYDROGEN TEST

See Section 3.1.2.2.1

3.1.3.5 DIARRHOEA

See Section 3.1.1.12

3.1.3.5.1 QUESTIONNAIRE

See Section 3.1.1.12.1

3.1.3.6 BLOATING

See Section 3.1.1.3

3.1.3.6.1 QUESTIONNAIRE

See Section 3.1.1.3.1

3.1.3.7 BORBORYGMI

See Section 3.1.1.5

3.1.3.7.1 QUESTIONNAIRE

See Section 3.1.1.5.1

3.1.3.8 FLATULENCE

See Section 3.1.1.8

3.1.3.8.1 QUESTIONNAIRE

See Section 3.1.1.8.1

3.1.3.9 ABDOMINAL PAIN/CRAMPS

See Section 3.1.1.2

3.1.3.9.1 VISUAL ANALOGUE SCALE

See Section 3.1.1.2.1

3.1.3.9.2 QUESTIONNAIRE

See Section 3.1.1.2.2

3.1.3.10 FAECAL BACTERIAL MASS

Faeces, the final product of GI activity, is composed of unfermented fibre, salts, water and bacteria. The number of bacteria (mostly anaerobes) in human faeces, estimated from direct microscopic counts, is between 1011 and 1012 per gram of dry faeces. It is estimated that 25% of wet stool weight and 50%-70% of dry stool weight (bacteria are about 80% water), come from bacterial mass, and that dietary fibre acts as a substrate for this mass (Stephen and Cummings 1980).

To evaluate the appropriateness of faecal bacterial mass as OV of maintenance of normal defecation,

the literature deriving from database #18 was critically evaluated (Table 1).

The large number of bacteria in stools indicates that bacterial growth has a dominating effect on total stool output. One of the factors influencing bacterial growth is diet. A major role of dietary component, in particular fibre, is to provide a substrate for fermentation by the microflora in the colon (Forsum et al. 1990). The result is to stimulate microbial growth and a greater excretion of microbial products in faeces. This leads to an increase in bacterial mass and consequently faecal mass, thus having a stool bulking effect. Increased bulk in the colon due to microbial proliferation decreases transit time. Furthermore, the presence of a high number of bacteria in faeces leads to an increase in gas production (carbon dioxide, hydrogen and methane) trapped in stool resulting in an increase in faecal bulk. The bulking effect induces a decrease in transit time.

Faecal bacterial mass does not represent a parameter directly correlated with the maintenance of normal defecation, but modifications leading to changes in variables, such as stool weight or transit time (bowel habits), may represent a problem for the maintenance of normal defecation, when compared to those of a normal situation.

In conclusion, the measurement of faecal bacterial mass, is not an appropriate parameter to be used for the substantiation of health claims in the context of maintenance of normal defecation. However, it can be used to support the postulated mechanisms by which the food/food component exerts the claimed effect.

3.1.3.10.1 GRAVIMETRIC PROCEDURE

Stephen and Cummings in 1980 have developed an accurate method to assess faecal bacterial mass, named as gravimetric procedure (Stephen and Cummings 1980). This method consists in separating the microbial fraction from the other faecal material, through the fractioning of faeces into three main components: bacteria, undigested fibre and soluble substances. Then, these fractions are weighted. The procedure has been developed from techniques used to isolate microbial matter from the rumen, with several altered (initial stomaching and filtering procedures in the presence of detergent) or omitted steps to improve the separation of bacteria from fibrous debris and to ensure the purity of the

bacterial fraction (Hoogenraad and Hird 1970). By this method, it is possible to obtain a direct estimate of the microbial contribution to the weight of the stool. The validation of effectiveness of the fractionation scheme was conducted in several studies by monitoring the location of muramic acid, an amino sugar found only in bacteria and conducting numerous bacterial counts, using stains specific for plant material, and measuring neutral sugars in wheat bran fibre. However, this method is time consuming because repeated washings and centrifugations are necessary to ensure a good separation of bacteria from other structural material in the stool.

In conclusion, gravimetric procedure is an appropriate method to assess faecal bacterial mass.

3.1.3.11 COMPOSITION OF THE GUT MICROBIOTA/ BIFIDOBACTERIAL POPULATION

See Section 3.1.1.15

3.1.3.11.1 16S rRNA MICROBIAL PROFILING

See Section 3.1.1.15.1

3.1.3.11.2 BIFIDOBACTERIAL ITS PROFILING

See Section 3.1.1.15.2

3.1.3.12 QUALITY OF LIFE

See Section 3.1.1.14

3.1.3.12.1 FUNCTIONAL DIGESTIVE DISORDERS QUALITY OF LIFE QUESTIONNAIRE

See Section 3.1.1.14.1

3.1.3.13 CONSTIPATION

See Section 3.1.1.10

3.1.3.13.1 PATIENT ASSESSMENT OF CONSTIPATION

See Section 3.1.1.10.1

3.1.3.14 SENSATION OF COMPLETE/INCOMPLETE EVACUATION

See Section 3.1.1.6

3.1.3.14.1 QUESTIONNAIRE

See Section 3.1.1.6.1

3.1.4 IMPROVING IRON ABSORPTION

3.1.4.1 NON-HAEM IRON ABSORPTION

Iron is a mineral naturally present in many foods and can be added to some food products, or used as a dietary supplement, inasmuch an adequate iron intake is essential for good health. In fact, iron is required for the functioning of proteins, such as haemoglobin (60%), myoglobin (5%), and for various enzymes involved in immune system functioning (5%). The remaining iron is found in body storage as ferritin (20%) and hemosiderin (10%), whereas a minor quantity (<0.1 %) is found as a transit chelate with transferrin. Dietary iron is present in two forms: as inorganic iron (ferrous and ferric compounds or non-haem iron) or organic forms (haem iron). Its availability is altered by many aspects, such as diet-related factors, including chemical forms of the nutrient, the type of cooking and processing of food, the presence of enhancers and inhibitors of iron absorption, as well as host-related factors like life-stage, nutritional and health status (Wienk et al. 1999). The inorganic iron is the predominant form of iron from vegetables and accounts for 80–90% of the iron in a standard diet, with the remaining 10% as haem iron. The latter derives primarily from haemoglobin and myoglobin, thus it is mainly associated with meat intake.

The iron balance is primarily regulated by controlling iron absorption and an imbalance of this mineral leads to nutritional deficiency or overload. Iron deficiency is the single most prevalent nutritional deficiency worldwide and leads to anaemia (http://www.who.int/nutrition/topics/ida/en/). Symptoms frequently associated with anaemia include pallor, weakness, fatigue, dyspnea, palpitations, sensitivity to cold, oral cavity and gastrointestinal tract abnormalities, and reduced capacity for work. In case of overload, iron is toxic and it is able to catalyse the formation of ROS.

To evaluate the appropriateness of non-haem iron absorption as OV of improving iron absorption, the literature deriving from database #19 was critically evaluated (Table 1).

Body iron concentration is kept within defined limits through precise mechanisms governing the

regulation of iron homeostasis; in particular, the iron amount in the body is determined by the regulation of iron absorption in the proximal small intestine.

Despite its relative scarcity, haem iron is absorbed far more efficiently than non-haem iron and may contribute up to 50% of the iron that actually enters the body. In fact, the bioavailability of ferrous iron (Fe^{2+}) is somewhat higher than that of ferric iron (Fe^{3+}), but haem iron is more efficiently absorbed than non-haem iron (Wienk et al. 1999). The amount of non-haem iron is strongly regulated by the intestinal mucosa (ferritin and then transferrin) to help assure that the total body amount of iron is within an acceptable range. In contrast, haem iron absorption is not strongly regulated and its absorption is not limited by the iron absorption control mechanism of the intestine. However, it is generally accepted that only soluble iron can be absorbed (Abbaspour et al. 2014). Soluble iron can be either in the ferric or in the ferrous form (non- haem iron), and it explains why all studies regarding iron solubility deal with non-haem iron.

In conclusion, the evaluation of non-haem iron absorption is an appropriate outcome variable for the substantiation of health claims in the context of improving iron absorption.

3.1.4.1.1 DOUBLE ISOTOPE TECHNIQUE

The determination of the amount of dietary mineral absorbed and retained by consuming diets characterized by different intakes, represents a valid approach in order to assess their human requirements. Several methods can be employed for this purpose, including, radioactive, stable isotope techniques or measurements using native iron.

Double isotope technique can be performed using both radioisotope or stable isotope (Kastenmayer et al. 1994). This technique can be obtained by injecting one isotope (^{55}Fe radioisotope or ^{58}Fe stable isotope) intravenously and giving the other (^{59}Fe radioisotope or ^{57}Fe stable isotope) orally, at the same time. The first isotope is used to determine the percentage of plasma iron used for haemoglobin synthesis. The isotopes are administered on consecutive days and enrichment of erythrocyte haemoglobin is measured 14 days after administration by transmutating stable isotope to radioisotopes by neutron-activation analysis, or directly by mass spectrometry (if stable isotope are

used) or by electroplating (for radioisotope).

Corrections for the natural abundance of the stable isotope have to be always performed. The use of two isotopes allows for correction of variations in iron clearance. Moreover, this method was validated against a well-accepted radioisotope and whole body counting method even though limited by the cost of the isotopes and the detection equipment.

Mainly in studies to perform in children and pregnant women, it is preferable to apply stable isotope techniques, owing to the advantages provided in comparison to other methods. Among these, it is possible to highlight their relatively more safety because of the lack of radioactive wastes.

In conclusion, double isotope technique represents an appropriate method to assess iron absorption.

3.1.4.1.2 WHOLE BODY COUNTING

Whole-body counting is a direct and possibly the most reliable measure for iron retention (Price et al. 1962). In this method ^{59}Fe (radioisotope that emits γ-rays) is given by mouth, and shortly afterwards the amount given is determined by external whole-body counting of radioactivity. After 10 to 14 days, when unabsorbed iron has been excreted, the amount of iron retained is determined by a further external whole-body measurement. Whole-body counting has the disadvantage of causing radiation exposure. Furthermore, the apparatus is expensive and the patient has to attend daily for counting. However, owing to its relative simplicity and repeatability, it is generally accepted as the reference method for iron absorption (Fairweather-Tait 2001). However, in studies to perform in children and pregnant women, it is preferable to apply methods that use stable isotope techniques.

In conclusion, whole body counting represents an appropriate method to assess iron absorption.

3.1.5 IMPROVING OF LACTOSE DIGESTION

3.1.5.1 BREATH HYDROGEN COPNCENTRATION

See Section 3.1.2.2

3.1.5.1.1 BREATH HYDROGEN TEST

See Section 3.1.2.2.1

3.1.5.2 NAUSEA

Nausea is an unpleasant symptom associated with different types of diseases and particular life conditions. Several causes lead to nausea (Linklater 2014) and generally they related to gastrointestinal (i.e. gastroparesis, gastric distension, and constipation), blood-borne (drugs and toxins) and vestibular (disruption of the inner ear often initiated by motion) factors. In addition, physiological states like pregnancy, or other conditions (e. g. infections, migraine headaches, motion sickness, food poisoning, cancer chemotherapy or other medicines) are often accompanied by nausea. It is an uneasy feeling in the stomach often accompanied by vomiting. The sensation of nausea reduces the quality of life and, even if not painful, is a very uncomfortable feeling that is felt in the chest, upper abdomen, or back of the throat. In some cases, nausea can be considered a reflex with a protective function that helps the body in reducing the digestion and absorption of ingested poisons, toxins or other substances that may be harmful for the health. Nausea may occur in acute and short-lived forms or chronically depending on the pathogenesis. In the latter case, the nausea is to be considered debilitating. People most affected from nausea are females, non-smokers, and those with history of motion sickness or postoperative disorders (about 30% of cases).

To evaluate the appropriateness of nausea as OV of improving lactose digestion, the literature deriving from database #20 was critically evaluated (Table 1).

Patients suffering from food allergy or food intolerances may have nausea, a symptom frequently difficult to describe for people. Therefore, food plays an important pathophysiological and therapeutic role (dietetic therapy for reduce sensation of nausea) for this symptom (Welliver 2013).

The most common form of food intolerance is lactose intolerance, which can trigger nausea. This disorder is characterised by a malassimilation of lactose that is therefore processed by colonic bacteria resulting in gas production, which in turn induces gastrointestinal distension. As a result, osmotic pressure increases in the colon and it accumulates water, leading to gastrointestinal symptoms such as diarrhoea, flatulence and nausea (Grand and Montgomery 2008).

However, nausea is not always present in patients who suffer from lactose malabsorption. In fact, some studies report diarrhoea, borborygmi, abdominal pain and flatulence as the main symptoms in

these subjects, whereas nausea occurs in a low percentage of patients. In the meanwhile, nausea can be associated with other detrimental conditions such as, gastroparesis, during chemotherapy or after anaesthesia, alcohol use disorders and more. For these reasons, nausea is a poor predictor of lactose maldigestion (Welliver 2013).

In conclusion, nausea is not an appropriate outcome variable to be used alone for the substantiation of claims in the context of improved lactose digestion. However, the sensation of nausea can be used as supportive evidence for such health claims.

3.1.5.2.1 QUESTIONNAIRE

Nausea, being a subjective symptom, is difficult to describe, and for this reason a valid measure of nausea is necessary for its assessment. There are different questionnaires that are used for evaluating nausea, but most of them do not take into account the complexity of this symptom. One of the most used is a modified version of McGill Pain Questionnaire, the McGill Nausea Questionnaire, in which the intensity of nausea is quantified with a visual-analogue scale (VAS) and an overall nausea intensity estimated by physicians and nurses on the basis of the patients' experience of nausea (Melzack et al. 1985). This questionnaire evaluates the experience of nausea itself and not just its frequency, severity, and duration. Although it is used in most studies, it is necessary to use a questionnaire with adjectives specifically designed to measure nausea in order to separate it from other subjective experiences, such as pain. The Nausea Profile (NP) (Muth et al. 1996) is a questionnaire that characterizes multiple dimensions of nausea, not only from a gastrointestinal experience but also from the somatic and emotional domains. It consists of 17 questions that are divided into three dimensions: somatic, gastrointestinal and emotional distress. Patients rated each of their symptoms on a scale from 0 (not at all) to 9 (severe). A total score is obtained by averaging the sum of all 17 questions and separate somatic, GI and emotional scores are calculated by the sums of selected questions. NP allows researchers to scale the total nausea experienced, but also it is able to establish a nausea profile, thanks to an individual's score on each of the 3 dimensions of nausea. Validity, reliability and sensibility of NP are based on the responses of undergraduates.

In conclusion, NP questionnaire appears to be a reliable and appropriate technique for assessing nausea.

3.1.5.3 DIARRHOEA

See Section 3.1.1.12

3.1.5.3.1 QUESTIONNAIRE

See Section 3.1.1.12.1

3.1.5.4 ABDOMINAL PAIN/CRAMPS

See Section 3.1.1.2

3.1.5.4.1 VISUAL ANALOGUE SCALE

See Section 3.1.1.2.1

3.1.5.4.2 QUESTIONNAIRE

See Section 3.1.1.2.2

3.1.5.5 BLOATING

See Section 3.1.1.3

3.1.5.5.1 QUESTIONNAIRE

See Section 3.1.1.3.1

3.1.5.6 FLATULENCE

See Section 3.1.1.8

3.1.5.6.1 QUESTIONNIARE

See Section 3.1.1.8.1

3.2 CLAIMS REFERRING TO CHILDREN DEVELOPMENT AND HEALTH

3.2.1 REDUCTION OF GI DISCOMFORT

3.2.1.1 CRYING TIME AND FREQUENCY

Crying has physiologic and neurophysiologic utility. The crying typically starts in the first few weeks of life and ends by age 4-5 months. Babies survive thanks to their first cry, because this serves as an effective force in the reorganization of extra uterine cardiorespiratory function. After birth, crying is

controlled by physiologic needs, such as hunger, temperature change, desire for attention and discomfort. Infants communicate their need by crying (St James-Roberts 1989). Healthy children cry on average nearly 3 hours per day at 6 weeks of age with a peak occurring between 3 PM and 11 PM. To evaluate the appropriateness of crying time and frequency as OV of reduction of GI discomfort, the literature deriving from database #21 was critically evaluated (Table 1).

Unexplained and recurrent bouts of crying in young children are often traditionally attributed to GI disturbances and discomfort/pain (Hyman et al. 2006). In particular, the term infant colic is commonly used to reflect this situation in infants. Infant colic is defined as an unexplained crying (excluding other reasons such as hungry, temperature change or desire of attention) of the otherwise healthy infant more than 3 hours a day and 3 days a week for at least 3 weeks and it was included in the list of childhood functional GI disorders of the Rome III Coordinating Committee. In addition, dyschezia is a GI disorder characterized by time of crying. In fact, it is defined as straining and crying for at least 10 min before successful passage of soft stools in an infant younger than 6 months of age without any other health problem.

In conclusion, evaluation of crying pattern appears an appropriate parameter for the substantiation of claims in the context of reduction of GI discomfort provided that other reasons for crying are excluded.

3.2.1.1.1 PARENTS' DIARY

The help of parents in reporting and interpreting symptoms is needed to assess time and frequency of crying. Parents' diary is the most widely used tool in studying crying patterns (Barr et al. 1988). A prospective assessment method is more reliable than retrospective one, because the latter is prone to recall bias. A validated 24-hour diary (study group was represented by 6 week old infants), developed by Barr et al. in the 1988, is the best diagnostic method to evaluate crying pattern (frequency and duration). The diary is composed by four 'time rulers' each representing six hours and vertical lines indicate five minute intervals. The rules must to be filled using symbols representing six behaviour patterns: sleeping, awake and content, fussing, crying, feeding, and sucking. Episodes of crying for

less than one minute are marked above the time rulers. In addition, parents must mark the type of feeding and the time of bowel movements.

Keeping a diary for 24 hours for seven or more days requires a high degree of parent co-operation. In particular, parents from lower social classes are less likely to participate or return diaries in survey studies and it seems impossible for parents to use this method daily for 12–16 weeks. However, as a compromise, it is possible to use this method during one predetermined day each week. Despite some limitations, these diaries may provide valid and useful reports of crying in the short term.

In conclusion, parents' diary appears an appropriate method for assess crying time and frequency.

3.2.1.2 ABDOMINAL DISTENSION

See Section 3.1.1.7

3.2.1.2.1 PARENTS' DARY

Most estimations of morbidity experienced by children are based on parental interviews/ questionnaires or on parental diaries, because it is necessary the help of parents for their reports and interpretations of symptoms, in particular when it is necessary assess subjective symptoms (Self et al. 2015). In addition, diaries can be useful in examining health event data when there is a need to monitor changes in vary symptoms in children. In general, a prospective assessment method (diary) is more reliable than retrospective one (interview or questionnaire), because the latter is prone to recall bias. However, diaries have intrinsic problems among which cost (mainly due to the method used to retrieve the diary records from respondents), respondent cooperation (non-perfect diary respondents tend to be younger adults, divorced/separated or never married, low-income, and low-educated) and diary completion. In addition to what previously said, diary is a more labor-intensive data collection method but it provides information about symptoms in children that may be impossible to collect by asking a parent to rely on their memory. Despite some limitations, parents' diaries may provide valid and useful reports.

In conclusion, parents' diary appears an appropriate method to assess diarrhoea, abdominal distention and pain, stool frequency and stool weight and constipation in children.

3.2.1.2.2 PARENTAL INTERVIEW

Diseases, discomfort and morbidity in children are assessed on the basis of parental interviews or diaries. The reliability and validity of these methods are difficult to evaluate and there are limitations in both. For parental interview, the main limitation is the telescoping effect. It refers to the temporal displacement of an event: recent events are recalled as happened earlier (backword telescoping) while remote events are perceived as happened more recently (forward telescoping) (Gaskell et al. 2000). In fact, parents tended to over-report events in retrospective data collection methods (parental interview) compared to prospective method (diary or medical records), underreporting occurred as well. Compared to the diary, the use of interview is recommended for low-grade education individuals due to the chance to have questions explained by the trained personnel. Finally, with the interview, it is possible recorded trivial symptoms that with the diary are lost. Although the use of parental interview for assessment of different children diseases or discomfort is widespread in field science, currently there are not sufficiently validated interviews for this purpose, although in several cases is the only method used to date for this purpose.

In conclusion, parental interview appears an appropriate method for the assessment of abdominal distention.

3.2.1.3 ABDOMINAL PIAN/CRAMPS

See Section 3.1.1.2

3.2.1.3.1 PARENTS' DIARY

See Section 3.2.1.2.1

3.2.1.4 DIARRHOEA

See Section 3.1.1.12

3.2.1.4.1 PARENTS' DIARY

See Section 3.2.1.2.1

3.2.1.5 CONSTIPATION

See Section 3.1.1.10

1541 3.2.1.5.1 PARENTS' DIARY

1542 See Section 3.2.1.2.1

1543 3.2.1.6 BREATH HYDROGEN CONCENTRATION

1544 See Section 3.1.2.2

1545 3.2.1.6.1 BREATH HYDROGEN TEST

1546 See Section 3.1.2.2.1

1547 3.2.1.7 INTESTINAL GAS VOLUME

1548 See Section 3.1.2.1

1549 3.2.1.7.1 MAGNETIC RESONANCE IMAGING

1550 See Section 3.1.2.1.1

1551 3.2.2 CONTRIBUTE TO SOFTENING OF STOOLS

1552 3.2.2.1 STOOL CONSISTENCY

1553 See Section 3.1.1.11

1554 3.2.2.1.1 BRISTOL STOOL SCALE

1555 See Section 3.1.1.11.1

1556 3.2.2.2 STOOL FREQUENCY

1557 See Section 3.1.1.13

1558 3.2.2.2.1 PARENTS' DIARY

1559 See Section 3.2.1.2.1

1560 3.2.2.3 STOOL WEIGHT/VOLUME/SIZE

1561 See Section 3.1.3.3

1562 3.2.2.3.1 DIRECT ASSESSMENT BY THE INVESTIGATORS

1563 See Section 3.1.3.3.1

1564 3.2.2.4 STOOL COLOUR

1565 The colour of children stools changes with age. In the early infancy, yellow is predominant in

1566 breastfed infants, whereas green coloured stools are occasionally reported in formula-fed infants,

probably because of the iron content of the formula (den Hertog et al. 2012). By 6 months the commonest stool colour tends to the brown and only in some occasions appears yellow or green. Black stools are uncommon at all ages (except for meconium), although they can be associated to an elevated iron content or to other dietary factors. Other possible stool colours are red and white. In this case they do not reflect a physiological condition but can be considered as a symptom due for example by GI bleeding or a liver dysfunction, respectively (Bekkali et al. 2009).

To evaluate the appropriateness of stool colour as OV of contribute to softening of stools, the literature deriving from database #11 was critically evaluated (Table 1).

Studies regarding infants ranging from one to three months of age pointed out a significant positive correlation between stool consistency and stool colour, independently of the type of feeding. For example, more lightly coloured stools (i.e. yellow) has been associated to increased fluidity of the stools. Other studies showed how the increased of the brown colour during the children life is probably related to the introduction of solids, which in turn increases stool consistency. However, despite these considerations, the stool colour can vary due to several factors, which might not influence the consistency (den Hertog et al. 2012). For example, red stools are caused by an infection, bleeding or colic polyps, whereas white stools can be a sign of a blockage in the liver. In addition, the intake of certain foods affects stool colour, in particular in children over three months when the diet start to vary. However, if it is considered only the range of "normal colour" (from yellow-brown, excluding white, red and black) and food and drink intakes are recorded, stool colour can be an appropriate parameter for evaluating softening of stool.

In conclusion, stool colour can be an appropriate outcome variable for the substantiation of health claims in the context of contribution to softening of stools.

3.2.2.4.1 PARENTS' DIARY

For evaluating stool colour in children, the "Amsterdam" Infant Stool Form Scale", which provides information concerning stool amount, consistency, and colour, has been developed (Bekkali et al. 2009). In order to classify the colour, 6 pictures illustrating the following colours are present: yellow,

orange, green, brown, meconium, and clay-coloured. Beside the colour category, also the categories of consistency and amount are present, each described in the scale by 4 photographs. The 14 pictures in total, are used as visual anchor points in this infant stool form scale. This scale can be used in daily standardized bowel diaries filled by parents. Despite this scale might be helpful in differentiating between normal and abnormal defecation patterns in infants, future studies are needed to validate its applicability and validity for research purposes.

In conclusion, parent's diary as the only one method used for assessing stool colour, is considered appropriate for this purpose.

3.2.3 INCREASE IN CALCIUM ABSORPTION

3.2.3.1 BONE MINERAL CONTENT

Bone mineral content (BMC) is a measurement of bone mineral found both in a specific area of the skeleton or in total skeleton system. Up to 50% by volume and 70% by weight of human bone is formed by hydroxyapatite, which is the mineral form of calcium apatite. BMC is expressed in grams (g) of hydroxyapatite and it is used to obtain bone mineral density (BMD), which is measured in grams per centimetre squared (g/cm^2), by dividing BMC by the area of the considered site (Ellis et al. 2001). Thus, due to the high association between BMD and BMC, it has been evidenced that also BMC is characterized by a growing phase during the childhood, depending on the availability of calcium and phosphate, with the following achievement of BMC peak during the early adulthood. After reaching peak bone mass, the mineral deposition activity of osteoblasts and the resorption activity of osteoclasts are balanced, leading to a steady state of the total BMC. Then, during adulthood, a constant and progressive imbalance of neo-mineralization and bone resorption, with prevailing osteoclast activity, causes a loss of BMD, reflecting a diminished BMC with ageing. Progressive loss of BMC results in osteopenia and osteoporosis. BMC, together with BMD and bone size is widely used in clinical practice for the assessment of the normal growth and development of bone in children. Additionally, by the fact that bone growth depends on hydroxyapatite deposition, BMC reflects calcium bioavailability in human body.

To evaluate the appropriateness of BMC as OV of the increase in calcium absorption, the literature deriving from database #22 was critically evaluated (Table 1).

BMC measurement, with adjustments for changes in body mass and total bone size, is widely performed in clinical practice for the assessment of bone health and mineralization in children and in adolescents (Budek et al. 2007, Ellis et al. 2001). BMC depends on both the size and density of skeletal bone, and a difference in BMC may reflect a difference in either bone size or bone density. BMC is the preferred outcome variable over BMD because bone expansion and the increase in BMC occur at different rate during childhood. Consequently, BMD calculated as BMC/bone area is not an appropriate ratio to be used in growing children because it is influenced by bone size (Ellis et al. 2001). Instead, it is well-accepted that bone mineralization should be assessed in three steps: height for age, bone area for height, and BMC for bone area. In comparative studies, it is important to adapt BMC measurement for age and sex, in order to adjust the heterogeneity in terms of the age- and sex-specific maturation (Ellis et al. 2001). Thus, to combine measurement results for children of different ages and to account for the growth-related changes in BMC, z-scores for BMC-for-age and BMC-for-height were calculated based on the healthy reference sample. In addition, because hydroxyapatite is primarily composed of calcium, BMC evaluation is also a useful tool in calcium bioavailability studies, which also allows to analyse the association between dietary intake and bone development and metabolism (Budek et al. 2007).

In conclusion, BMC is an appropriate outcome variable for the scientific substantiation of health claims in the context of increase in calcium absorption in children.

3.2.3.1.1 DUAL ENERGY X-RAY ABSORPTIOMETRY

Dual Energy X- Ray Absorptiometry (DXA), also known as bone densitometry or bone density scanning, can accurately analyse bone and non-bone tissue, providing a quantification of BMD, BMC, fat mass and soft lean mass. It has been validated across age groups, from premature infants to older adults, including both normal and overweight subjects. The use of DXA in infants and children is gradually increasing, with the aim to understand the impact of disease on bone health or nutritional

impact on body composition. Indeed, DXA is also a useful tool for assessing the whole skeletal maturity, the body fat composition. Moreover, it is used for evaluating the efficacy of pharmaceutical therapy. DXA is a peculiar imaging modality which differs from other X-ray systems because requires special beam filtering and near perfect spatial registration of two attenuations. Indeed, DXA system creates a two-dimension image resulting from the combination of low and high energy attenuations. Although density is typically given by mass per volume unit, DXA can only quantify the bone density as a mass per area unit, since it uses planar images and cannot measure the bone depth. By the fact that a two-dimensional output is given, DXA-based bone mass cannot distinguish between bone compartments, namely cortical and trabecular bone (Nilsson 2015). For this reasons DXA measurement can be integrated with additional 3D outputs from different technologies, as quantitative computed tomography (QCT). Nevertheless, it is regarded as safe, with a minimal radiation exposure (0.1 μGy), relatively fast (6-7 min for total body assessment) and highly reproducible (Deng et al. 2002). On the other hand, DEXA is expensive and requires specific skills. Whole body DXA scans is primarily used for BMC measurements in children (Budek et al. 2007) and for body composition measurements in adults, while several common measurement sites, including the lumbar spine, the proximal hip and the forearm, are preferred when measuring BMD. For the set-up of RCTs, DXA measurement should

be performed at baseline and then not earlier than 12 months, which is considered the most appropriate follow-up interval to detect (if any) significant changes in BMD and/or BMC.

In summary, DXA is generally an appropriate method to assess BMD and BMC, in human intervention studies.

3.2.3.1.2 SINGLE PHOTON ABSORPTIOMETRY

In the early 1960s, a new method for bone densitometry, called single photon absorptiometry (SPA), was developed to overcome the problems of previous radiographic photodensitometric techniques caused by polychromatic X-rays and non-uniform film sensitivity. Indeed, SPA technique uses a single energy gamma ray source (^{125}I) photon energy, and a scintillation detector to measure the

single-energy photon beam passage through bone and soft tissue. The distal radius (wrist) is usually used as the site of measurement because the amount of soft tissue in this area is small. Changes in beam intensity are due to the attenuation by bone mineral and the integrated attenuation is proportional to the mass of mineral in the scan path, whose length is proportional to the width of the bone. Even if SPA has been widely used in the past for the assessment of bone mineral density and content (Neer 1992), it is outdated and nowadays it has been replaced by other densitometry techniques, such as Dual Photon Absorptiometry and DXA which have greater accuracy and are capable of measuring central skeletal sites. In fact, the radionuclide source (^{125}I) emitted an average energy of 27 keV, which is sufficient for the BMC measurement of appendicular bones but not for that of central skeletal sites. Other limitations are represented by the use of radionuclides, which gradually decays and requires regular replacement, and by the scanning time (15-30 minutes), which is considerable because of the low rate of photon flux. With the low scanning, undesirable drawbacks might occur, such as the patient moving during the scan leading to poor quality of the scan image and so limiting the reproducibility. Moreover, SPA method can compensate for variation in bone width but not for variation in bone thickness. The reproducibility of the measurement therefore depends upon the ability to reproduce exactly the location of the measurement. For this reason, it is necessary to control the stillness and the pronation/supination of the bone site (generally the forearm), since rotation alters the photon beam path (Neer 1992).

In summary, even if it was a widely used bone densitometric technique, SPA is not an appropriate method for assess BMC.

3.2.3.2 BONE MINERAL DENSITY

Bone mass is considered a synonym of BMD; indeed, based on the evaluation methodology, bone mass accounts for the sum of two components: areal BMD, which is a two-dimensional measurement, expressed in g/cm^2, usually obtained through DXA scans, and volumetric BMD, expressed in g/cm^3, which is a 3D measure given by QCT; volumetric BMD can discriminate between cortical and trabecular bone, thus emerging as qualitative, other than quantitative medical tool only.

Physiologically, BMD reaches its peak in the early adulthood both in males and females and subsequently declines with ages from the fifth decade (Rizzoli 2014), even if lifestyle (e.g. cigarette smoking, excessive alcohol consumption, prolonged immobilization) or genetic factors can accelerate this process. On the opposite, bone mass increases in response to increased mechanical stimuli (e.g. physical activity and gravity), that are able to at least maintain bone homeostasis. Bone mass is also influenced by ethnic differences and sex (Curtis et al. 2015).

To evaluate the appropriateness of BMD as OV of the increase in calcium absorption, the literature deriving from database #22 was critically evaluated (Table 1).

Bone is a composite tissue made up of an organic collagen protein and inorganic mineral (hydroxyapatite). Bone mineral density (BMD, g/cm^2 or BMC/bone area), is a measure of bone density and consequently, it provides an estimate of stored calcium in bone tissue. However, if BMD is used to compare bone of different size and thickness differences, it can be incorrectly interpreted (Carter et al. 1992). Furthermore, an important factor to be taken into account for the assessment of BMD is that BMC not always correlates to bone area. This is because their relationship depends on different factors, including the population group, the body size, the skeletal site, as well as the instrumental and scanning conditions (Prentice et al. 1994). This may lead to erroneous results regarding other size-related variables of bones such as calcium intake. In particular in children BMC is the preferred outcome variable over BMD because bone expansion and the increase in BMC occur at different rate during childhood. Consequently, BMD calculated as BMC/bone area is not an appropriate ratio to be used in growing children because it is influenced by bone size.

BMD values are expressed as T and Z scores. In adult, the World Health Organization (WHO) criterion for diagnosing osteoporosis is based on BMD T scores, defined as the standard deviation (SD) score of the observed BMD compared with that of a normal young adult. However, due to the above mentioned reasons, T scores are not appropriate for growing children and should not be used. The use of the Z score, defined as the SD score based on age-specific and sex- specific norms, is considered a more appropriate method of comparison of BMD in children. If the Z score is below -

2.0, the International Society of Clinical Densitometry recommends the use of the terminology “low bone density for chronological age” (Lewiecki et al. 2004).

In conclusion, BMD is not an appropriate parameter for the scientific substantiation of health claims in the context of increase in calcium absorption in children.

3.2.3.2.1 DUAL ENERGY X-RAY ABSORPTIOMETRY

See Section 3.2.3.1.1

3.2.3.3 CALCIUM BALANCE

Calcium balance is generally defined as the difference between the dietary intake and the output (faecal and urinary) of Ca, the most abundant mineral in the human body. Consequently, it can be positive, negative or neutral. Ca is involved in several physiological functions, including bone growth, nerve conduction, muscle contraction and blood coagulation. Approximately 99% of total body Ca is contained in bones, whereas the remaining fraction is within extracellular fluids and soft tissue (Hsu and Levine 2004). Calcium metabolism is affected by parathyroid hormone, 1,25-dihydroxy-vitamin D (1,25-D) and calcitonin. These three hormones act together in order to maintain serum Ca concentration at nearly constant values, directly conditioning intestinal Ca absorption, renal re-absorption, Ca excretion and utilization of Ca in the bone (Bass and Chan 2006).

To evaluate the appropriateness of calcium balance as OV of the increase in calcium absorption, the literature deriving from database #23 was critically evaluated (Table 1).

A negative Ca balance, determined in presence of output exceeding input, represents a state leading over the time to its depletion that contributes to skeletal demineralization. On the contrary, a positive balance is associated with an accrual and repletion of Ca stores, contributing to the maintenance of bone health. Alterations in calcium metabolism, observed as chronic hyper- or hypocalcaemia, may lead to serious clinical problems. The former may predispose to vascular calcifications and nephrocalcinosis, whereas the latter, relatively more common in children, in conjunction with deficiencies of vitamin D, may result in rickets or osteomalacia, with a major impact on health, growth and development of infants, children and adolescent (Allgrove 2003).

The measurement of whole-body Ca balance is affected by some aspects making its assessment challenging. It could be skewed by erroneous determination of faecal Ca losses, which affect the results more than incorrect calculations of urinary losses. This can be explained by the need to collect faeces over a period of 5-10 days to be representative of the diet. Consequently, faecal Ca losses are up to 10 times greater than urinary losses. Among dietary factors, besides Ca, phosphorous and protein intake exert an influence on urinary Ca excretion, potentially modulating Ca balance (Calvez et al. 2012). Thus, in order to improve the interpretation of data obtained, the net absorption of Ca should be measured, distinguishing into the faecal output the unabsorbed dietary amount of Ca and the amount secreted into the intestine and not reabsorbed (usually referred to as endogenous faecal excretion). Furthermore, especially for children < 4 years old, there is an absence of data, mainly due to the impracticality of prolonged dietary regulation and complete urine and faecal collections that are required for traditional balance studies especially in children who are not toilet-trained.

In conclusion, calcium balance alone is not an appropriate outcome variable for substantiation of health claims regarding increase in calcium absorption in children.

3.2.3.3.1 STABLE ISOTOPE TECHNIQUES

The determination of the amount of dietary mineral absorbed and retained by consuming diets characterized by different intakes, represents a valid approach in order to assess their human requirements. Several methods can be employed for this purpose, including mass-balance measurements, radioactive or stable isotope techniques (Abrams 1999).

Owing to the presence of six stable isotopes of Ca with different natural distribution, this mineral is particularly adequate for studies with isotopes, now more available and less expensive.

Mainly in studies to perform in children, it is preferable to apply stable isotope techniques, owing to the advantages provided in comparison to other methods. Among these, it is possible to highlight:

- Their relatively more safety because of the lack of radioactive wastes. Their adaptability to longitudinal studies performed in order to evaluate the modulation of growth and development on dietary Ca requirements.

- Their ability to distinguish, from the faecal output, both the amount of unabsorbed dietary Ca and the amount of endogenous faecal excretion. These two sources of Ca in the faeces are not provided by mass balance studies (Abrams 1999).

Ca absorption can be calculated using different isotopic methods. Among these, single-isotopic technique involves an isotope of the mineral ingested either with a meal or separately. The collection of faeces is completed when virtually the entire unabsorbed isotope is recovered. The difference between the amount ingested and recovered in the faeces represents the fraction of tracer absorbed. This method provides the benefit of calculating only dietary Ca, without including endogenous secretory losses. At the other end of the spectrum, a long period of faeces collection is required.

More information (e.g. endogenous faecal Ca excretion) can be obtained by dual tracer technique that applies a low-abundance stable isotope ingested and a different-one injected intravenously. After administration of the tracers, a complete 24-h urine collection is carried out. The amount of oral isotope absorbed is represented by the relative fraction, in the 24-h urine pool, of the ingested isotope compared with the intravenous amount (Abrams 1999). Although spot determinations of urine or serum isotope concentrations may also be employed, this method is less accurate than 24-h collection(Yergey et al. 1994). In order to assess endogenous faecal excretion of Ca, the injection of a large dose of the tracer and a collection of faeces for a period of 6-7 days (3-4 in infants) are necessary.

The determination of isotopic content of blood, urine and faecal samples can be obtained using different methodologies, such as irradiation and mass spectrometry. The former, first-developed, is relatively cumbersome compared to the latter.

In conclusion, stable isotope techniques are appropriate methods of measurement of Ca balance.

3.2.4 IMPROVING IRON ABSORPTION

3.2.4.1 NON- HAEM IRON ABSORPTION

See Section 3.1.4.1

3.2.4.1.1 DOUBLE ISOTOPE TECHNIQUE

1801 See Section 3.1.4.1.1

1802 3.2.4.1.2 WHOLE BODY COUNTING

1803

1804 **4. CONCLUSIONS**

1805 Several foods and food components have been the object of applications for authorization of health
1806 claims pursuant to Regulation (EC) 1924/2006. Most of them have received a negative for many
1807 reasons, including the choice of not appropriate OVs and/or MMs. The present manuscript provides
1808 information related to the collection, collation and critical analysis of claimed effects, OVs and MMs
1809 that have been proposed so far in the context of GI health, compliant with the European Regulation.
1810 This work could help EFSA to develop further guidance to applicants in the preparation of new
1811 applications for authorization of health claims in the context of oral health.
1812 Moreover, this critical evaluation may help stakeholder with interest in requesting the authorization
1813 for the use of a health claim related to GI health. Despite many aspects (e.g. adequate sample size,
1814 study design and statistical analysis) are crucial for receiving a positive opinion from EFSA, this
1815 work may indeed help during the choice of OVs and MMs to be considered in human intervention
1816 studies.
1817 In addition to the use for health claim substantiation, this critical evaluation of OVs and MMs can
1818 impact general research, being useful for the design of human intervention studies, independently
1819 from health claim substantiation.

1820

1821 **CONFLICT OF INTEREST**

1822 All authors have no conflicts of interest

1823

1824 **ACKNOWLEDGMENTS:**

This project has received financial support from the European Food Safety Authority (EFSA), Grant GP/EFSA/NUTRI/2014/01. The present article, however, is under the sole responsibility of the authors. The positions and opinions presented in this article are those of the authors alone and do not necessarily represent the views/any official position or scientific works of EFSA. To know about EFSA guidance documents and other scientific outputs of EFSA, please consult its website at: http://www.efsa.europa.eu.

1833 **REFERENCES**

1834 Abbaspour N, Hurrell R, Kelishadi R. 2014. Review on iron and its importance for human health. J
1835 Res Med Sci. 19:164-174.
1836 Abrams SA. 1999. Using stable isotopes to assess mineral absorption and utilization by children. Am
1837 J Clin Nutr. 70:955-964.
1838 Agachan F, Chen T, Pfeifer J, Reissman P, Wexner SD. 1996. A constipation scoring system to
1839 simplify evaluation and management of constipated patients. Dis Colon Rectum. 39:681-685.
1840 Agrawal A, Whorwell PJ. 2008. Review article: abdominal bloating and distension in functional
1841 gastrointestinal disorders--epidemiology and exploration of possible mechanisms. Aliment
1842 Pharmacol Ther. 27:2-10.
1843 Alame AM, Bahna H. 2012. Evaluation of constipation. Clin Colon Rectal Surg. 25:5-11.
1844 Allen SJ, Okoko B, Martinez E, Gregorio G, Dans LF. 2004. Probiotics for treating infectious
1845 diarrhoea. Cochrane Database Syst Rev.CD003048.
1846 Allgrove J. 2003. Disorders of calcium metabolism. Current Paediatrics. 13:529-535.
1847 Arce DA, Ermocilla CA, Costa H. 2002. Evaluation of constipation. Am Fam Physician. 65:2283-
1848 2290.
1849 Azpiroz F, Guyonnet D, Donazzolo Y, Gendre D, Tanguy J, Guarner F. 2015. Digestive Symptoms
1850 in Healthy People and Subjects With Irritable Bowel Syndrome: Validation of Symptom
1851 Frequency Questionnaire. J Clin Gastroenterol. 49:e64-70.
1852 Azpiroz F, Malagelada JR. 2005. Abdominal bloating. Gastroenterology. 129:1060-1078.
1853 Bannister JJ, Davison P, Timms JM, Gibbons C, Read NW. 1987. Effect of stool size and consistency
1854 on defecation. Gut. 28:1246-1250.
1855 Barr RG, Kramer MS, Boisjoly C, Mcvey-White L, Pless IB. 1988. Parental diary of infant cry and
1856 fuss behaviour. Arch Dis Child. 63:380-387.
1857 Basilisco G, De Marco E, Tomba C, Cesana BM. 2007. Bowel urgency in patients with irritable bowel
1858 syndrome. Gastroenterology. 132:38-44.

1859 Bass JK, Chan GM. 2006. Calcium nutrition and metabolism during infancy. Nutrition. 22:1057-
1860 1066.
1861 Bekkali N, Hamers SL, Reitsma JB, Van Toledo L, Benninga MA. 2009. Infant stool form scale:
1862 development and results. J Pediatr. 154:521-526 e521.
1863 Belsey J, Greenfield S, Candy D, Geraint M. 2010. Systematic review: impact of constipation on
1864 quality of life in adults and children. Aliment Pharmacol Ther. 31:938-949.
1865 Borgaonkar MR, Irvine EJ. 2000. Quality of life measurement in gastrointestinal and liver disorders.
1866 Gut. 47:444-454.
1867 Budek AZ, Hoppe C, Ingstrup H, Michaelsen KF, Bugel S, Molgaard C. 2007. Dietary protein intake
1868 and bone mineral content in adolescents-The Copenhagen Cohort Study. Osteoporos Int.
1869 18:1661-1667.
1870 Calvez J, Poupin N, Chesneau C, Lassale C, Tome D. 2012. Protein intake, calcium balance and
1871 health consequences. Eur J Clin Nutr. 66:281-295.
1872 Carter DR, Bouxsein ML, Marcus R. 1992. New approaches for interpreting projected bone
1873 densitometry data. J Bone Miner Res. 7:137-145.
1874 Chang L, Lee OY, Naliboff B, Schmulson M, Mayer EA. 2001. Sensation of bloating and visible
1875 abdominal distension in patients with irritable bowel syndrome. Am J Gastroenterol. 96:3341-
1876 3347.
1877 Chassany O, Marquis P, Scherrer B, Read NW, Finger T, Bergmann JF, Fraitag B, Geneve J, Caulin
1878 C. 1999. Validation of a specific quality of life questionnaire for functional digestive disorders.
1879 Gut. 44:527-533.
1880 Collins SM. 2014. A role for the gut microbiota in IBS. Nat Rev Gastroenterol Hepatol. 11:497-505.
1881 Cummings JH (2001). The Effect of Dietary Fiber on Fecal Weight and Composition. *In:* Spiller G
1882 A (ed.) CRC Handbook on dietary fiber in nutrition. Boca Raton, FL: CRC Press.
1883 Curtis E, Litwic A, Cooper C, Dennison E. 2015. Determinants of Muscle and Bone Aging. J Cell
1884 Physiol. 230:2618-2625.

1885 Da Fonseca EG (2015). Quality of life factors in bladder and bowel dysfunction. *In:* FRANCO, I.,
1886 Austin PF, Bauer SB, Von Gontard A, Homsy Y. (eds.) Pediatric incontinence. Chichester, UK:
1887 John Wiley & Sons, Ltd.
1888 Davies GJ, Crowder M, Reid B, Dickerson JW. 1986. Bowel function measurements of individuals
1889 with different eating patterns. Gut. 27:164-169.
1890 Den Hertog J, Van Leengoed E, Kolk F, Van Den Broek L, Kramer E, Bakker EJ, Bakker-Van Gijssel
1891 E, Bulk A, Kneepkens F, Benninga MA. 2012. The defecation pattern of healthy term infants up
1892 to the age of 3 months. Arch Dis Child Fetal Neonatal Ed. 97:F465-470.
1893 Deng HW, Xu FH, Davies KM, Heaney R, Recker RR. 2002. Differences in bone mineral density,
1894 bone mineral content, and bone areal size in fracturing and non-fracturing women, and their
1895 interrelationships at the spine and hip. J Bone Miner Metab. 20:358-366.
1896 Dimidi E, Christodoulides S, Fragkos KC, Scott SM, Whelan K. 2014. The effect of probiotics on
1897 functional constipation in adults: a systematic review and meta-analysis of randomized
1898 controlled trials. Am J Clin Nutr. 100:1075-1084.
1899 EFSA. 2016. Guidance on the scientific requirements for health claims related to the immune system,
1900 the gastrointestinal tract and defence against pathogenic microorganisms EFSA J. 14:4369.
1901 Ellis KJ, Shypailo RJ, Hardin DS, Perez MD, Motil KJ, Wong WW, Abrams SA. 2001. Z score
1902 prediction model for assessment of bone mineral content in pediatric diseases. J Bone Miner Res.
1903 16:1658-1664.
1904 Fairweather-Tait SJ. 2001. Iron. J Nutr. 131:1383S-1386S.
1905 Felce D, Perry J. 1995. Quality of life: its definition and measurement. Res Dev Disabil. 16:51-74.
1906 Forsum E, Eriksson C, Goranzon H, Sohlstrom A. 1990. Composition of faeces from human subjects
1907 consuming diets based on conventional foods containing different kinds and amounts of dietary
1908 fibre. Br J Nutr. 64:171-186.
1909 Francis CY, Morris J, Whorwell PJ. 1997. The irritable bowel severity scoring system: a simple
1910 method of monitoring irritable bowel syndrome and its progress. Aliment Pharmacol Ther.

11:395-402.

Frank L, Flynn J, Rothman M. 2001. Use of a self-report constipation questionnaire with older adults in long-term care. Gerontologist. 41:778-786.

Frank L, Kleinman L, Farup C, Taylor L, Miner P, Jr. 1999. Psychometric validation of a constipation symptom assessment questionnaire. Scand J Gastroenterol. 34:870-877.

Gallagher EJ, Liebman M, Bijur PE. 2001. Prospective validation of clinically important changes in pain severity measured on a visual analog scale. Ann Emerg Med. 38:633-638.

Gareau MG, Sherman PM, Walker WA. 2010. Probiotics and the gut microbiota in intestinal health and disease. Nat Rev Gastroenterol Hepatol. 7:503-514.

Gaskell GD, Wright DB, O'muircheartaigh CA. 2000. Telescoping of landmark events: implications for survey research. Public Opin Q. 64:77-89.

Ghoshal UC, Gupta D, Kumar A, Misra A. 2007. Colonic transit study by radio-opaque markers to investigate constipation: validation of a new protocol for a population with rapid gut transit. Natl Med J India. 20:225-229.

Grand RJ, Montgomery RK. 2008. Lactose malabsorption. Curr Treat Options Gastroenterol. 11:19-25.

Groll D, Vanner SJ, Depew WT, Dacosta LR, Simon JB, Groll A, Roblin N, Paterson WG. 2002. The IBS-36: a new quality of life measure for irritable bowel syndrome. Am J Gastroenterol. 97:962-971.

Halmos EP, Power VA, Shepherd SJ, Gibson PR, Muir JG. 2014. A diet low in FODMAPs reduces symptoms of irritable bowel syndrome. Gastroenterology. 146:67-75 e65.

Hawker GA, Mian S, Kendzerska T, French M. 2011. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken). 63 Suppl 11:S240-252.

1937 Hayes V, Morris J, Wolfe C, Morgan M. 1995. The SF-36 health survey questionnaire: is it suitable
1938 for use with older adults? Age Ageing. 24:120-125.
1939 Hays RD, Morales LS. 2001. The RAND-36 measure of health-related quality of life. Ann Med.
1940 33:350-357.
1941 Heaton KW, Cripps HA. 1993. Straining at stool and laxative taking in an English population. Dig
1942 Dis Sci. 38:1004-1008.
1943 Heaton KW, Radvan J, Cripps H, Mountford RA, Braddon FE, Hughes AO. 1992. Defecation
1944 frequency and timing, and stool form in the general population: a prospective study. Gut. 33:818-
1945 824.
1946 Hoogenraad NJ, Hird FJ. 1970. The chemical composition of rumen bacteria and cell walls from
1947 rumen bacteria. Br J Nutr. 24:119-127.
1948 Houghton LA. 2011. Bloating in constipation: relevance of intraluminal gas handling. Best Pract Res
1949 Clin Gastroenterol. 25:141-150.
1950 Hsu SC, Levine MA. 2004. Perinatal calcium metabolism: physiology and pathophysiology. Semin
1951 Neonatol. 9:23-36.
1952 Hyman PE, Milla PJ, Benninga MA, Davidson GP, Fleisher DF, Taminiau J. 2006. Childhood
1953 functional gastrointestinal disorders: neonate/toddler. Gastroenterology. 130:1519-1526.
1954 Iovino P, Bucci C, Tremolaterra F, Santonicola A, Chiarioni G. 2014. Bloating and functional gastro-
1955 intestinal disorders: where are we and where are we going? World J Gastroenterol. 20:14407-
1956 14419.
1957 Irvine EJ, Tack J, Crowell MD, Gwee KA, Ke M, Schmulson MJ, Whitehead WE, Spiegel B. 2016.
1958 Design of Treatment Trials for Functional Gastrointestinal Disorders. Gastroenterology.
1959 150:1469-1480 e1461.
1960 Irvine EJ, Whitehead WE, Chey WD, Matsueda K, Shaw M, Talley NJ, Veldhuyzen Van Zanten SJ.
1961 2006. Design of treatment trials for functional gastrointestinal disorders. Gastroenterology.
1962 130:1538-1551.

1963 Jellema P, Schellevis FG, Van Der Windt DA, Kneepkens CM, Van Der Horst HE. 2010. Lactose
1964 malabsorption and intolerance: a systematic review on the diagnostic value of gastrointestinal
1965 symptoms and self-reported milk intolerance. QJM. 103:555-572.
1966 Jenkinson C, Coulter A, Wright L. 1993. Short form 36 (SF36) health survey questionnaire:
1967 normative data for adults of working age. BMJ. 306:1437-1440.
1968 Kastenmayer P, Davidsson L, Galan P, Cherouvrier F, Hercberg S, Hurrell RF. 1994. A double stable
1969 isotope technique for measuring iron absorption in infants. Br J Nutr. 71:411-424.
1970 Knowles CH, Aziz Q. 2009. Basic and clinical aspects of gastrointestinal pain. Pain. 141:191-209.
1971 Kusano M, Hosaka H, Kawada A, Kuribayashi S, Shimoyama Y, Zai H, Kawamura O, Yamada M.
1972 2014. Gastrointestinal motility and functional gastrointestinal diseases. Curr Pharm Des.
1973 20:2775-2782.
1974 Lackner JM, Jaccard J, Keefer L, Firth R, Carosella AM, Sitrin M, Brenner D. 2014. The accuracy of
1975 patient-reported measures for GI symptoms: a comparison of real time and retrospective reports.
1976 Neurogastroenterol Motil. 26:1802-1811.
1977 Lam C, Chaddock G, Marciani Laurea L, Costigan C, Cox E, Hoad C, Pritchard S, Gowland P, Spiller
1978 R. 2017. Distinct Abnormalities of Small Bowel and Regional Colonic Volumes in Subtypes of
1979 Irritable Bowel Syndrome Revealed by MRI. Am J Gastroenterol. 112:346-355.
1980 Lane MM, Czyzewski DI, Chumpitazi BP, Shulman RJ. 2011. Reliability and validity of a modified
1981 Bristol Stool Form Scale for children. J Pediatr. 159:437-441 e431.
1982 Lee KS, Kang DS, Yu J, Chang YP, Park WS. 2012. How to do in persistent diarrhea of children?:
1983 concepts and treatments of chronic diarrhea. Pediatr Gastroenterol Hepatol Nutr. 15:229-236.
1984 Leibbrand R, Cuntz U, Hiller W. 2002. Assessment of functional gastrointestinal disorders using the
1985 Gastro-Questionnaire. Int J Behav Med. 9:155-172.
1986 Lewiecki EM, Watts NB, Mcclung MR, Petak SM, Bachrach LK, Shepherd JA, Downs RW, Jr. 2004.
1987 Official positions of the international society for clinical densitometry. J Clin Endocrinol Metab.
1988 89:3651-3655.

1989 Li LF, Chan RL, Lu L, Shen J, Zhang L, Wu WK, Wang L, Hu T, Li MX, Cho CH. 2014. Cigarette
1990 smoking and gastrointestinal diseases: the causal relationship and underlying molecular
1991 mechanisms (review). Int J Mol Med. 34:372-380.
1992 Linklater GT. 2014. Nausea and vomiting. J R Coll Physicians Edinb. 44:218-223.
1993 Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, Spiller RC. 2006. Functional
1994 bowel disorders. Gastroenterology. 130:1480-1491.
1995 Malinen E, Rinttila T, Kajander K, Matto J, Kassinen A, Krogius L, Saarela M, Korpela R, Palva A.
1996 2005. Analysis of the fecal microbiota of irritable bowel syndrome patients and healthy controls
1997 with real-time PCR. Am J Gastroenterol. 100:373-382.
1998 Manichanh C, Eck A, Varela E, Roca J, Clemente JC, Gonzalez A, Knights D, Knight R, Estrella S,
1999 Hernandez C, Guyonnet D, Accarino A, Santos J, Malagelada JR, Guarner F, Azpiroz F. 2014.
2000 Anal gas evacuation and colonic microbiota in patients with flatulence: effect of diet. Gut.
2001 63:401-408.
2002 Marquis P, De La Loge C, Dubois D, Mcdermott A, Chassany O. 2005. Development and validation
2003 of the Patient Assessment of Constipation Quality of Life questionnaire. Scand J Gastroenterol.
2004 40:540-551.
2005 Martinez-Martinez MI, Calabuig-Tolsa R, Cauli O. 2017. The effect of probiotics as a treatment for
2006 constipation in elderly people: A systematic review. Arch Gerontol Geriatr. 71:142-149.
2007 Martinez AP, De Azevedo GR. 2012. The Bristol Stool Form Scale: its translation to Portuguese,
2008 cultural adaptation and validation. Rev Lat Am Enfermagem. 20:583-589.
2009 Martini D, Angelino D, Cortelazzi C, Zavaroni I, Bedogni G, Musci M, Pruneti C, Passeri G, Ventura
2010 M, Galli D, Mirandola P, Vitale M, Dei Cas A, Bonadonna RC, Di Nuzzo S, De Felici MB, Del
2011 Rio D. 2018a. Claimed Effects, Outcome Variables and Methods of Measurement for Health
2012 Claims Proposed Under European Community Regulation 1924/2006 in the Framework of
2013 Maintenance of Skin Function. Nutrients. 10:7.
2014 Martini D, Biasini B, Rossi S, Zavaroni I, Bedogni G, Musci M, Pruneti C, Passeri G, Ventura M,

2015 Galli D, Mirandola P, Vitale M, Dei Cas A, Bonadonna RC, Del Rio D. 2017a. Claimed effects,
2016 outcome variables and methods of measurement for health claims on foods proposed under
2017 European Community Regulation 1924/2006 in the area of appetite ratings and weight
2018 management. Int J Food Sci Nutr. *doi: 10.1080/09637486.2017.1366433.*
2019 Martini D, Guareschi C, Biasini B, Bedogni G, Galli C, Angelino D, Marchi L, Zavaroni I, Pruneti
2020 C, Ventura M, Galli D, Mirandola P, Vitale M, Dei Cas A, Bonadonna RC, Passeri G, Del Rio
2021 D. 2018b. Claimed effects, outcome variables and methods of measurement for health claims
2022 proposed under Regulation (EC) 1924/2006 in the framework of bone health. PharmaNutrition.
2023 6:17-36.
2024 Martini D, Rossi S, Biasini B, Zavaroni I, Bedogni G, Musci M, Pruneti C, Passeri G, Ventura M, Di
2025 Nuzzo S, Galli D, Mirandola P, Vitale M, Dei Cas A, Bonadonna RC, Del Rio D. 2017b. Claimed
2026 effects, outcome variables and methods of measurement for health claims proposed under
2027 European Community Regulation 1924/2006 in the framework of protection against oxidative
2028 damage and cardiovascular health. Nutr Metab Cardiovasc Dis. 27:473-503.
2029 Mathur R, Amichai M, Chua KS, Mirocha J, Barlow GM, Pimentel M. 2013. Methane and hydrogen
2030 positivity on breath test is associated with greater body mass index and body fat. J Clin
2031 Endocrinol Metab. 98:E698-702.
2032 Mccoll E. 2004. Best practice in symptom assessment: a review. Gut. 53 Suppl 4:iv49-54.
2033 Melzack R. 1975. The McGill Pain Questionnaire: major properties and scoring methods. Pain. 1:277-
2034 299.
2035 Melzack R, Rosberger Z, Hollingsworth ML, Thirlwell M. 1985. New approaches to measuring
2036 nausea. CMAJ. 133:755-758, 761.
2037 Milani C, Hevia A, Foroni E, Duranti S, Turroni F, Lugli GA, Sanchez B, Martin R, Gueimonde M,
2038 Van Sinderen D, Margolles A, Ventura M. 2013. Assessing the fecal microbiota: an optimized
2039 ion torrent 16S rRNA gene-based analysis protocol. PLoS One. 8:e68739.
2040 Milani C, Lugli GA, Turroni F, Mancabelli L, Duranti S, Viappiani A, Mangifesta M, Segata N, Van

Sinderen D, Ventura M. 2014. Evaluation of bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing (ITS) protocol. FEMS Microbiol Ecol. 90:493-503.

Muller-Lissner S, Koch G, Talley NJ, Drossman D, Rueegg P, Dunger-Baldauf C, Lefkowitz M. 2003. Subject's Global Assessment of Relief: an appropriate method to assess the impact of treatment on irritable bowel syndrome-related symptoms in clinical trials. J Clin Epidemiol. 56:310-316.

Murakami K, Okubo H, Sasaki S. 2006. Dietary intake in relation to self-reported constipation among Japanese women aged 18-20 years. Eur J Clin Nutr. 60:650-657.

Muth ER, Stern RM, Thayer JF, Koch KL. 1996. Assessment of the multiple dimensions of nausea: the Nausea Profile (NP). J Psychosom Res. 40:511-520.

Myo K, Thein Win N, Kyaw-Hla S, Thein Thein M, Bolin TD. 1994. A prospective study on defecation frequency, stool weight, and consistency. Arch Dis Child. 71:311-313; discussion 313-314.

Neer RM. 1992. The utility of single-photon absorptiometry and dual-energy x-ray absorptiometry. J Nucl Med. 33:170-171.

Nilsson AG. 2015. Bone research: an issue of maturity. J Intern Med. 277:626-629.

North American Society for Pediatric Gastroenterology HaN. 2006. Evaluation and treatment of constipation in children: summary of updated recommendations of the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition. J Pediatr Gastroenterol Nutr. 43:405-407.

Pares D, Comas M, Dorcaratto D, Araujo MI, Vial M, Bohle B, Pera M, Grande L. 2009. Adaptation and validation of the Bristol scale stool form translated into the Spanish language among health professionals and patients. Rev Esp Enferm Dig. 101:312-316.

Perman JA, Modler S, Barr RG, Rosenthal P. 1984. Fasting breath hydrogen concentration: normal values and clinical application. Gastroenterology. 87:1358-1363.

2067 Peters HP, De Vries WR, Vanberge-Henegouwen GP, Akkermans LM. 2001. Potential benefits and
2068 hazards of physical activity and exercise on the gastrointestinal tract. Gut. 48:435-439.
2069 Prentice A, Parsons TJ, Cole TJ. 1994. Uncritical use of bone mineral density in absorptiometry may
2070 lead to size-related artifacts in the identification of bone mineral determinants. Am J Clin Nutr.
2071 60:837-842.
2072 Price DC, Cohn SH, Wasserman LR, Reizenstein PG, Cronkite EP. 1962. The determination of iron
2073 absorption and loss by whole body counting. Blood. 20:517-531.
2074 Price KR, Lewis J, Wyatt GM, Fenwick GR. 1988. Flatulence--causes, relation to diet and remedies.
2075 Nahrung. 32:609-626.
2076 Pritchard SE, Marciani L, Garsed KC, Hoad CL, Thongborisute W, Roberts E, Gowland PA, Spiller
2077 RC. 2014. Fasting and postprandial volumes of the undisturbed colon: normal values and changes
2078 in diarrhea-predominant irritable bowel syndrome measured using serial MRI.
2079 Neurogastroenterol Motil. 26:124-130.
2080 Rana SV, Malik A. 2014. Hydrogen breath tests in gastrointestinal diseases. Indian J Clin Biochem.
2081 29:398-405.
2082 Rao DR, Bello H, Warren AP, Brown GE. 1994. Prevalence of lactose maldigestion. Influence and
2083 interaction of age, race, and sex. Dig Dis Sci. 39:1519-1524.
2084 Rasquin-Weber A, Hyman PE, Cucchiara S, Fleisher DR, Hyams JS, Milla PJ, Staiano A. 1999.
2085 Childhood functional gastrointestinal disorders. Gut. 45 Suppl 2:II60-68.
2086 Rey E, Balboa A, Mearin F. 2014. Chronic constipation, irritable bowel syndrome with constipation
2087 and constipation with pain/discomfort: similarities and differences. Am J Gastroenterol. 109:876-
2088 884.
2089 Rizzoli R. 2014. Nutritional aspects of bone health. Best Pract Res Clin Endocrinol Metab. 28:795-
2090 808.
2091 Rose C, Parker A, Jefferson B, Cartmell E. 2015. The Characterization of Feces and Urine: A Review
2092 of the Literature to Inform Advanced Treatment Technology. Crit Rev Environ Sci Technol.

45:1827-1879.

Sandler RS, Jordan MC, Shelton BJ. 1990. Demographic and dietary determinants of constipation in the US population. Am J Public Health. 80:185-189.

Sanjoaquin MA, Appleby PN, Spencer EA, Key TJ. 2004. Nutrition and lifestyle in relation to bowel movement frequency: a cross-sectional study of 20630 men and women in EPIC-Oxford. Public Health Nutr. 7:77-83.

Self MM, Williams AE, Czyzewski DI, Weidler EM, Shulman RJ. 2015. Agreement between Prospective Diary Data and Retrospective Questionnaire Report of Abdominal Pain and Stooling Symptoms in Children with Irritable Bowel Syndrome. Neurogastroenterol Motil. 27:1110-1119.

Spiegel BM, Bolus R, Agarwal N, Sayuk G, Harris LA, Lucak S, Esrailian E, Chey WD, Lembo A, Karsan H, Tillisch K, Talley J, Chang L. 2010. Measuring symptoms in the irritable bowel syndrome: development of a framework for clinical trials. Aliment Pharmacol Ther. 32:1275-1291.

Spiller RC. 1994. Intestinal absorptive function. Gut. 35:S5-9.

St James-Roberts I. 1989. Persistent crying in infancy. J Child Psychol Psychiatry. 30:189-195.

Stephen AM, Cummings JH. 1980. The microbial contribution to human faecal mass. J Med Microbiol. 13:45-56.

Stevens J, Vansoest PJ, Robertson JB, Levitsky DA. 1987. Mean transit time measurement by analysis of a single stool after ingestion of multicolored plastic pellets. Am J Clin Nutr. 46:1048-1054.

Stewart WF, Liberman JN, Sandler RS, Woods MS, Stemhagen A, Chee E, Lipton RB, Farup CE. 1999. Epidemiology of constipation (EPOC) study in the United States: relation of clinical subtypes to sociodemographic features. Am J Gastroenterol. 94:3530-3540.

Suarez FL, Levitt MD. 2000. An understanding of excessive intestinal gas. Curr Gastroenterol Rep. 2:413-419.

Sullivan SN. 2012. Functional abdominal bloating with distention. ISRN Gastroenterol.

2012:721820.

Svedlund J, Sjodin I, Dotevall G. 1988. GSRS--a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci. 33:129-134.

Tack J, Janssen P. 2010. Gastroduodenal motility. Curr Opin Gastroenterol. 26:647-655.

Talley NJ, Boyce PM, Owen BK, Newman P, Paterson KJ. 1995. Initial validation of a bowel symptom questionnaire and measurement of chronic gastrointestinal symptoms in Australians. Aust N Z J Med. 25:302-308.

Tomlin J, Lowis C, Read NW. 1991. Investigation of normal flatus production in healthy volunteers. Gut. 32:665-669.

Van Der Sijp JR, Kamm MA, Nightingale JM, Britton KE, Mather SJ, Morris GP, Akkermans LM, Lennard-Jones JE. 1993. Radioisotope determination of regional colonic transit in severe constipation: comparison with radio opaque markers. Gut. 34:402-408.

Viniol A, Keunecke C, Biroga T, Stadje R, Dornieden K, Bosner S, Donner-Banzhoff N, Haasenritter J, Becker A. 2014. Studies of the symptom abdominal pain--a systematic review and meta-analysis. Fam Pract. 31:517-529.

Wald A, Sigurdsson L. 2011. Quality of life in children and adults with constipation. Best Pract Res Clin Gastroenterol. 25:19-27.

Weaver LT, Steiner H. 1984. The bowel habit of young children. Arch Dis Child. 59:649-652.

Welliver M. 2013. Nausea and vomiting: mechanisms and treatment overview. Gastroenterol Nurs. 36:378-380.

Wienk KJ, Marx JJ, Beynen AC. 1999. The concept of iron bioavailability and its assessment. Eur J Nutr. 38:51-75.

Williamson A, Hoggart B. 2005. Pain: a review of three commonly used pain rating scales. J Clin Nurs. 14:798-804.

Wong RK, Drossman DA. 2010. Quality of life measures in irritable bowel syndrome. Expert Rev Gastroenterol Hepatol. 4:277-284.

2145 Yergey AL, Abrams SA, Vieira NE, Aldroubi A, Marini J, Sidbury JB. 1994. Determination of
2146 fractional absorption of dietary calcium in humans. J Nutr. 124:674-682.

2147

2148