REVIEW ARTICLE OPEN

Techniques and Methods

The mixed-meal tolerance test as an appetite assay: methodological and practical considerations

James A. King[1,2✉], Alice E. Thackray[1,2], Catherine Gibbons[3], Catia Martins[4], David R. Broom[5], David J. Stensel[1,2,6,7], Dimitris Papamargaritis[2,8], Frank Arsenyadis[2,8], Graham Finlayson[3], Gráinne Whelehan[2,8], Javier T. Gonzalez[9], John Blundell[3], Kristine Beaulieu[3], Lewis James[1,2], Lore Metz[10], Mark Hopkins[11], Masashi Miyashita[1,6,7], Scott A. Willis[1,2], Vicky Drapeau[12] and David Thivel[10]

© The Author(s) 2025

Appetite control is a topic which attracts widespread interest given its importance to energy balance and obesity. In this research area, the mixed-meal tolerance test (MM-TT) has emerged as an '*appetite regulation assay*', facilitating the dynamic assessment of appetite parameters (e.g. subjective appetite perceptions, appetite-related hormones, food reward) in response to an individual meal. The MM-TT is commonly employed in observational and experimental studies to examine population differences and intervention effects. Problematically, no practice standard exists for the MM-TT and protocols vary widely. This presents a challenge for researchers designing new MM-TTs and hampers the comparability of findings. Therefore, within this narrative review we sought to identify and discuss key methodological considerations inherent within a MM-TT. The scope of our review extends to evaluating participant familiarisation and methodological standardisation practices, test meal characteristics, appetite perception assessment, blood sampling techniques, measurement of appetite-related hormones and data handling/analysis. A checklist has been devised to summarise relevant methodological issues identified within this review. This checklist can be used as a tool by researchers to facilitate MM-TT design and promote greater standardisation/comparability between studies. This review highlights the need for broader standardisation of MM-TT procedures to support consistency across future research. Additional research is needed to strengthen the evidence base on which various recommendations are made, particularly relating to participant familiarisation and methodological standardisation practices. Additional scrutiny of less common outcomes employed in MM-TTs (not addressed here), such as diet-induced thermogenesis, gastric emptying and *ad libitum* energy intake, is also needed.

International Journal of Obesity (2025) 49:2168–2183; https://doi.org/10.1038/s41366-025-01866-7

INTRODUCTION

Appetite control is a topic which has risen in prominence, spurred largely (but not exclusively) by the need to tackle global obesity. This interest emanates from basic scientists keen to understand the psychobiological foundations of appetite [1, 2], as well as applied scientists focused on interventions (lifestyle, pharmaceutical, surgical). With the multi-dimensional nature of appetite, one of the key challenges in this field is the issue of measurement [3, 4]. That is, to identify which appetite constructs should be assessed and how their measurement is operationalised. This involves consideration of accuracy, precision and validity of measures [3] alongside a clear theoretical multi-disciplinary framework in which these constructs are considered.

Within appetite and energy balance science, the mixed-meal tolerance test (MM-TT) is a research method frequently employed as an '*appetite regulation assay*'. Practically, this method involves the repeated assessment of psychological and biological appetite parameters before and after a meal challenge (Fig. 1). The mixed (macronutrient) nature of this stimulus facilitates ecological validity and elicits holistic meal-related responses (cognitive and metabolic). The primary outcomes typically include appetite perceptions (i.e. visual analogues scales [VAS] for hunger, fullness, satisfaction & desire-to-eat) and circulating appetite-related

[1]National Centre for Sport and Exercise Medicine, School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, UK. [2]NIHR Leicester Biomedical Research Centre, University Hospitals of Leicester NHS Trust and University of Leicester, Leicester, UK. [3]School of Psychology, Faculty of Medicine & Health, University of Leeds, Leeds, UK. [4]Department of Nutrition Sciences, The University of Alabama at Birmingham, Birmingham, AL, USA. [5]Centre for Physical Activity, Sport and Exercise Sciences, Coventry University, Coventry, UK. [6]Faculty of Sport Sciences, Waseda University, Saitama, Japan. [7]Department of Sports Science & Physical Education, The Chinese University of Hong Kong, Hong Kong, China. [8]Diabetes Research Centre, University of Leicester, Leicester, UK. [9]Centre for Nutrition, Exercise and Metabolism, Department of Health, University of Bath, Bath, UK. [10]Clermont Auvergne University, EA 3533, Laboratory of the Metabolic Adaptations to Exercise under Physiological and Pathological Conditions (AME2P), CRNH, Clermont-Ferrand, France. [11]School of Food Science & Nutrition, Faculty of Environment, University of Leeds, Leeds, UK. [12]Department of Kinesiology, Faculty of Medicine, Université Laval, Quebec City, QC, Canada. ✉email: j.a.king@lboro.ac.uk

Received: 24 April 2025 Revised: 20 June 2025 Accepted: 15 July 2025
Published online: 30 July 2025

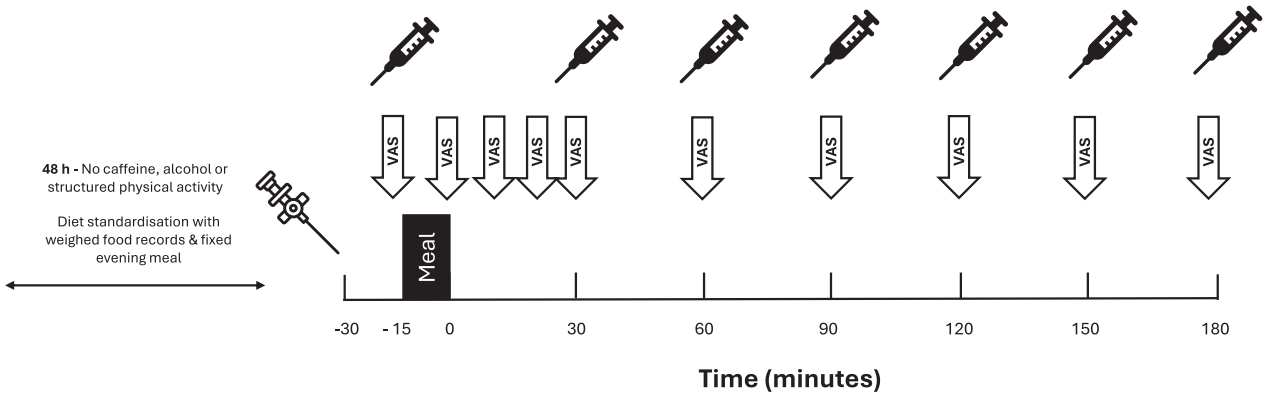


Fig. 1 An example MM-TT. Intravenous cannulation is performed 30 min before the MM-TT for 'habituation.' Blood samples are collected for the analysis of appetite-related peptides. Each participant's daily energy requirements are calculated based on their resting metabolic rate multiplied by their physical activity level (PAL). A balanced meal (50% carbohydrate, 35% fat, 15% protein) is consumed within 15 min after which postprandial measures begin. Note that a more frequent and longer duration blood sampling regimen may be desirable depending on the research question. VAS, 100 mm visual analogue scales for hunger, fullness, desire-to-eat & prospective food consumption.

hormone concentrations (e.g. ghrelin, peptide-YY [PYY], glucagon-like peptide-1 [GLP-1]). Additional outcomes also include food reward (e.g. Leeds Food Preference Questionnaire) [5], substrate oxidation, diet-induced thermogenesis and gastric emptying. The MM-TT is typically performed in a controlled laboratory environment with prior standardisation of dietary intake and lifestyle behaviours. A key principle of the MM-TT is that the dynamic assessment of appetite responses to a meal provides greater insight into appetite regulation than isolated, static (often fasted) assessments. The latter assessments typically capture a single time-point and may not reflect the complex, integrated physiological and psychological responses to eating that occur across the day. The protocol has theoretical parallels with the satiety cascade conceptual framework, originally proposed by Blundell et al. [6]. Specifically, the satiety cascade describes the timing and sequence of cognitive and physiological events that occur in response to food consumption which promote satiation (processes that bring an eating episode to an end) and satiety (processes that inhibit eating in the postprandial period). Herein, the MM-TT measures the post-ingestive consequences of eating food and therefore provides a measurement of satiety.

The MM-TT is commonly employed in observational (cross-sectional) studies to investigate differences between populations. For instance, comparisons between people with and without obesity [7, 8], or those who are physically active versus inactive [9]. The MM-TT is also used in experimental research to assess the effect of acute and chronic interventions. The latter include lifestyle interventions (e.g. exercise and diet) [10, 11], bariatric surgery [12] and obesity pharmacotherapy [13–15]. Additional applications include metabolic disease risk prediction in prospective studies [16, 17]. Problematically, there is no standard of practice for the MM-TT and protocols vary considerably across studies. This variation poses a challenge for researchers when designing new experiments and reduces comparability of findings. Methodological considerations include pre-test standardisation, test meal size and composition, test duration and measurement routine, assessment tools (e.g. questionnaires), blood sampling technique, analytes/biochemical assays, and data handling (statistical analysis) procedures.

This paper discusses key methodological elements about MM-TT design to support more informed decisions in human (predominantly adult) appetite research. The merits of different approaches are highlighted alongside recommendations for best practice. Because the MM-TT is a discrete technique, the scope of this review is limited to procedural elements that directly influence the MM-TT itself. Therefore, wider influences on energy homoeostasis or the longer-term control of appetite (e.g. body composition, physical activity, and energy expenditure) are not addressed. Equally, the use of MM-TTs to investigate other outcomes (e.g. metabolic and cardiovascular) is out of scope for this paper.

FAMILIARISATION AND STANDARDISATION

Familiarisation considerations

Ahead of MM-TTs, participants should be familiarised with research environments and phlebotomy procedures to attenuate anxiety and potential stress responses which may impact appetite [18, 19]. Once an intravenous cannula has been inserted, it is good practice to allow a 'habituation period' (15–30 min) before baseline assessments are made [18]. Participants should also be familiarised with study test meals (ideally consuming the foods) to attenuate novelty effects and VAS to ensure that participants understand scale constructs and descriptors.

Standardisation considerations—meal presentation and consumption

Appetite and eating behaviour are influenced by many psychobiological factors. It is therefore important that the physical (i.e. appearance) and verbal presentation of the test meal are standardised. Specifically, participants should not be subject to external food cues during MM-TTs. This includes food odours [20], experimenter interaction, and exposure via the internet/mobile phones, videos and literature. Research has shown that even the size of cutlery and crockery require standardisation [21].

As appetite-related outcomes are sensitive to circadian rhythms, MM-TTs should begin at a fixed time that is replicated within and/or between participants. Typically, MM-TTs should begin early in the morning so that participants do not have to fast excessively. Morning tests—as opposed to afternoon or evening—also avoid logistical and cost issues associated with providing standardised breakfast and lunch meals (necessary for later MM-TTs). The time for test meal consumption must also be fixed (often 10–15 min) and standardised within participants so that meal duration does not confound outcomes. Moreover, for multi-item meals, the order of consumption should be standardised. Importantly, researchers must ensure that all of the test meal is consumed so that the stimulus is received as intended. To ensure this is achievable during the selected time-frame, pilot testing and familiarisation may be necessary. During the MM-TT, water consumption should be limited to a sensible standardised volume with the timing of consumption also standardised. The temperature and humidity of

research environments should be comfortable (22 to 23 °C), monitored throughout the MM-TT, and replicated within and between participants. Environmental temperature is particularly important as it impacts appetite perceptions [22] and may arterialise blood, potentially impacting concentrations of circulating analytes.

Standardisation considerations—pre-test diet

Standardisation of diet is commonplace before MM-TTs, particularly when repeated assessments are made within participants. This practice helps avoid pre-test energy and nutritional imbalance. To facilitate this, MM-TTs often commence in the morning with participants fasted, having not eaten since the previous evening. The timing and content of this final meal are commonly fixed. Weighed food records are often employed to facilitate dietary replication in the 24–72 h before MM-TTs. When financially and practically feasible, research teams may provide full standardisation diets. One study showed that 48 h of dietary standardisation was associated with lower variation in hunger responses to repeated test meals compared with no standardisation [23]. Without more evidence, a pre-test standardisation period of at least 48 h is recommended.

Hydration and alcohol intake are additional dietary considerations. Prolonged hypohydration may induce haemoconcentration and impact appetite perceptions [24]. Voluntary water intake is highly variable between individuals, but including water (and other fluid) intake as a recorded variable in the pre-testing food diaries can help to control this, at least within-participants, before testing. To help ensure euhydration, ~35–40 ml/kg/day of water (from food and drinks) in the days leading up to testing facilitates this, as well as consumption of a water bolus (~500 mL) in the hours before MM-TTs. However, consumption shortly before the MM-TT (i.e. <30 min) should be avoided, as acute water intake may suppress appetite and *ad libitum* energy intake [25]. Alcohol consumption may acutely affect appetite, eating behaviour, substrate metabolism, gastro-intestinal function and fluid balance [26]. As alcohol may specifically influence substrate metabolism for up to 24 h [27], we recommend prohibition of alcohol for 48 h before MM-TTs [23].

Standardisation considerations—physical activity

Physical activity is commonly standardised in the 24–72 h before MM-TTs due to its influence on energy balance and acute appetite-modulating effects [28]. Additionally, participants should avoid physical exertion when travelling to research sites before MM-TTs. Accelerometry can be used to confirm physical activity standardisation in the days before MM-TTs (and during extended trials). While physical activity can be monitored in many ways, one potential method to facilitate standardisation is to pair a daily activity diary with any research-grade device capable of sensitively measuring steps per day. Device-measured steps provide an objective measure of overall physical activity volume and captures less formal or incidental movement. Concomitantly, an activity log facilitates the measurement of more formal activity, including that not captured by accelerometry (e.g. cycling) and the details of formal 'training.' Note that when studying active populations, researchers may wish to allow some activity in the days before MM-TTs—as the elimination of activity may represent an intervention. Recording and replication of activity is crucial in these scenarios. From a practical perspective, researchers may wish to limit any activity to that which is habitually undertaken by participants. Within the 24 h before the MM-TT, it may be prudent to request any activity to be completed early in the day to enable recovery. Without specific evidence to base a recommendation, a 48-h physical activity standardisation period may be prudent before MM-TTs. Some support for standardising physical activity 48 h before MM-TTs comes from evidence that exercise can improve insulin sensitivity within this time [29].

Standardisation considerations—sleep

Sleep restriction impacts appetite perceptions, appetite-related hormones, food preferences and reward [30–32]. For instance, higher circulating ghrelin concentrations and hunger ratings were seen after just one night of sleep restriction [30]. Research participants should therefore be instructed to maintain adequate (7–9 h) and consistent sleep routines in the days before MM-TTs. A sleep diary (paper record of sleep onset and wake time) and/or an accelerometer capable of monitoring sleep variables—particularly night-time sleep onset and morning wake times—can facilitate this. Within individuals, standardisation of wake time is particularly important on the morning of MM-TTs [33].

Standardisation considerations—smoking and caffeine

Smoking and caffeine intake are commonly controlled in appetite studies. As nicotine in tobacco suppresses appetite [34], tobacco/cigarette smokers are often excluded from appetite studies. If smokers are included, smoking routines should be standardised and recorded in the days before assessments. Participants should not be asked to stop smoking during this period, as withdrawal effects may affect study outcomes. On the day of MM-TTs, participants may smoke before MM-TTs (which must be replicated on successive assessments) but not during the procedure. Electronic cigarettes (e-cigarettes) and vapes also contain nicotine in combination with artificial flavourings and other chemicals which may alter the hedonic value and cravings for specific foods [35]. For these, the same standardisation procedures should be followed as with smoking [36, 37]. Marijuana (cannabis) may also influence appetite via orexigenic properties [38]. Unlike tobacco, only a small subset are daily or dependent users [39]; therefore, excluding these individuals may be appropriate unless central to the research question. In recreational users, although evidence is limited, withdrawals are uncommon and possible orexigenic effects are generally thought to persist directly and indirectly for up to 12 h [39–41]. Accordingly, abstinence for at least 12 h is recommended.

Caffeine may increase metabolic rate and suppress appetite and/or energy intake [42]. Therefore, highly caffeinated foods and beverages (especially coffee) are commonly prohibited 24–48 h before MM-TTs. However, the half-life of caffeine is ~4 h and withdrawal effects may be equally disruptive to behaviour. Consequently, moderation of intake, with standardisation across repeated tests, may be more appropriate in the prior 24–48 h; with abstinence on the day of the MM-TT. As the half-life of caffeine is ~73% longer (10.7 ± 3.0 h) in women taking oral contraceptives [43], researchers may wish to extend the abstinence period.

Standardisation considerations—menstrual phase and contraceptives

Self-reported energy intake varies across the menstrual cycle in eumenorrheic women, being higher in the luteal versus follicular phase [44]. Cyclical changes in circulating sex hormones, impacting appetite perceptions, food cravings and food reward, may underpin this finding [45]. Standardisation of the menstrual phase is therefore important when investigating women in appetite-related research [46]. In experimental research, all women should be tested in the same menstrual cycle phase. Furthermore, it is important to consider which cycle phase women are tested in—especially if research involves investigating sex-based differences. Most often, comparisons are made to women in the follicular phase of the menstrual cycle, because oestrogen and progesterone levels are low, stable and more similar to men. Testing in the follicular phase therefore helps to limit any confounding influence of sex hormones on behaviour and metabolism. However, it must be recognised that women's responses may differ if tested in the luteal phase, particularly as mood, cravings, water retention and gastrointestinal symptoms may differ. Therefore, testing in the luteal phase may provide a

more 'real-world' comparison. In either case, unless women are tested in both phases, limitations of the approach taken must be explicitly considered. The variability in menstrual hormones between women means that within-subject designs are preferred over between-subject designs.

Verification of menstrual cycle phase is essential [46]. Whilst detailed monitoring of individual sex hormone profiles is the gold standard [46], this technique may be unfeasible. Other techniques that can be used in combination include cycle counting (forward and backward methods), ovulation monitoring (i.e. luteinising hormone tests), body temperature measurement, and wearable technology (i.e. multi-sensor tracking of physiological signals associated with cycle phase) [46, 47]. Given the difficultly in determining cycle phase, it is recommended that women with irregular, anovulatory or hormonally abnormal cycles should be identified in screening, analysed separately or excluded from research—depending on the research question at hand [46].

Whether individuals are taking oral contraceptives or are using other forms of hormonal contraception (e.g. intrauterine device) is an additional consideration in studies making comparisons between people (and should be documented). Whilst evidence is limited, oral contraceptive use may impact appetite-related neuro-peptide metabolism [48]. Consequently, it has been suggested that people taking contraception should form separate study groups and not be mixed with samples of participants who are not taking contraception [46, 47].

Standardisation considerations—medications

A range of drug classes are known to influence appetite, energy intake, energy expenditure and/or body weight. These therapies may confound research when comparing different cohorts, or assessing intervention responses, particularly if medications change during a study. This is an important consideration when developing study inclusion/exclusion criteria and highlights the necessity to conduct thorough participant medication histories before enrolment. Whilst it is not possible to compile a complete list here, relevant common medications include certain anti-depressant and anti-anxiety medications [49], thyroid medications [50], corticosteroids [51] and therapies for obesity [52] (e.g. GLP-1-based medications) and diabetes (e.g. insulin, glitazones, metformin, SGLT2-inhibitors) [53].

MEAL STIMULUS

Meal energy content

Deciding on the energy content of the meal stimulus is a challenging issue when designing MM-TTs. Initially, researchers must decide whether meals are provided as a fixed amount for all participants (akin to an oral glucose tolerance test in diabetes medicine), or whether energy content is '*scaled*' to a participant-level parameter. The former option is simple but can introduce between- and within-participant response variation, especially when the sample is physically diverse or has undergone significant intervention. The latter is more frequently chosen but raises further considerations about how the meal is scaled. Historically, meal energy content has most commonly been normalised to body mass, i.e. a set number of kilocalories per kilogram of body mass (akin to how various expert societies set out energy intake recommendations). This approach benefits from simplicity but is problematic when research participants vary substantially regarding adiposity e.g. if a study included people with and without obesity. In such cases, scaling to fat-free mass may be better. Alternatively, when feasible, meals can be scaled to resting metabolic rate (RMR) [54] or total daily energy expenditure (TDEE). The latter requires an additional estimate of movement-related energy expenditure (mainly necessary if participants vary significantly in their habitual activity levels), most precisely measured via doubly-labelled water techniques; but more economically and practically feasible via an accelerometer/inclinometer capable of estimating energy expenditure [55, 56]. Equally, movement-related energy expenditure can be estimated via self-report techniques, for example, established Physical Activity Levels (PAL, TDEE = RMR x PAL) [57]. In these cases, meal energy content is prescribed as a percentage of an individual's measured or equation-derived (less optimal) values.

Within each scenario, researchers must determine the exact numerical values on which to scale. Practically, this calculation is made based on the test meal being scientifically and ecologically valid regarding size i.e. providing sufficient energy to elicit a meaningful appetite response (perceptually and metabolically) without invoking nausea or other adverse gastrointestinal symptoms, and similar to what participants would habitually eat (in a single meal context). When scaling test meals based on RMR and subsequent TDEE, meals are typically 20% to 30% of an individual's daily energy requirements, to clearly differentiate a meal from snacking [58]. While standardising meal content across studies enhances comparability, there may be specific instances where non-standard scaling may be necessary. For example, individuals with reduced gastric capacity (e.g. post-bariatric surgery) or gastrointestinal conditions (e.g. gastroparesis, inflammatory bowel disease, functional dyspepsia) may not tolerate standard meal sizes, necessitating smaller/modified meals [59, 60]. This may also influence the MM-TT duration (typically 2–4 h) where the assessment period length may be modified given altered gastric emptying or other gastrointestinal dysfunction [59, 61]. Nevertheless, within a study, it is important all participants receive the same meal stimulus for internal validity, and test meals are trialled prior to the study for acceptability/tolerability.

Macronutrient composition

Test meal macronutrient contribution is often determined as a secondary feature of the actual foods provided in a MM-TT. This is frequently determined based on palatability considerations for the local population (i.e. culturally-preferred foods) and practical considerations about meal provision (e.g. available preparation facilities). However, the composition of the test meal within a MM-TT is important as it impacts appetite and metabolic responses postprandially. For example, on a per kilocalorie basis, protein generates the strongest satiety response, and fat the weakest [62]. Moreover, the magnitude of response of appetite-related peptides to meals varies depending on the macronutrient stimulus [63]. Beyond this, qualitative differences in the types of carbohydrate, protein and fat, or the confounding effect of soluble/insoluble fibre, may further impact postprandial appetite and metabolic responses. For example, offering equal amounts of dairy-based casein versus whey protein, which have known differences in amino acid profile and postprandial metabolic response, have differing postprandial VAS scores for satiety and fullness [64]. Similarly, isoenergetic meals with higher or lower ratios of simple versus complex carbohydrates also provoke different postprandial effects on hunger, satiety and metabolic parameters [65]. Additionally, polyunsaturated and saturated fatty acids differentially impact post-meal appetite-related hormone responses [66]. Therefore, differences in test meal composition directly contribute to variability in meal responses between studies, which hampers comparability of research findings.

When considering MM-TTs using whole foods, it is practically impossible to recommend a universal standardised test meal to use given international differences in habitual dietary intake, food availability and food preferences. However, most studies in the literature have opted for '*balanced*' test meals containing ~50% carbohydrate, ~35% fat, and ~15% protein. This likely reflects Western food preferences and the predominance of research studies from high-income Western nations in the scientific literature. While specific food items selected may differ between studies based on cultural norms or studying populations with

specific dietary requirements (e.g. gluten intolerance), researchers may still wish to opt for a '*balanced*' profile. Such a meal profile is broadly consistent with international recommendations for healthy diets and ensures that each macronutrient is sufficiently represented within the meal stimulus. In the absence of a popular alternative, researchers may wish to replicate this meal profile to facilitate comparability.

Physical state

The physical state of the test meal is another consideration for researchers undertaking MM-TTs. Whilst most studies base their meal stimulus on whole foods, oral nutritional supplements, i.e. nutritionally-complete, ready-to-drink liquid meals (often milkshakes or juices), can also be used [13]. Whilst there are merits in each approach, the postprandial appetitive response is qualitatively different. Importantly, whilst metabolic and hormonal responses are comparable, liquid meals produce weaker satiety responses compared with solid meals [67]. Nonetheless, oral nutritional supplements benefit from their ease of preparation (reducing human error introduced while weighing and preparing food), homogeneity in content, and validated nutritional profile (including the ability to exclude allergens). Moreover, the nature of oral nutritional supplements facilitates swift consumption, potentially reducing variability in meal duration between people, and allows for blinding where relevant for certain research questions. Equally, participants' lack of familiarity with oral nutritional supplements attenuates the impact of prior interactions/preferences on satiation and satiety, which may be particularly important in cross-sectional studies involving single assessments. Alternatively, meals based on whole foods have greater ecological validity i.e. better reflect real-life. Notably, the requirement to chew whole foods influences a range of psychobiological (appetitive) responses [68], which are part of the normal process of eating. Overall, the approach taken will be influenced by the research question at hand alongside practical considerations.

APPETITE SCALES/QUESTIONNAIRES

Appetite-related self-reports include several items intended to capture, over a given period, specific subjective sensations or perceived motivations to eat. These include general feelings of hunger/fullness, desire to eat (in general or specific food types), and prospective judgements about the quantity of food or specific food types that could or would be eaten. Appetite-related VAS have been used extensively in studies examining motivation to eat in relation to food preloads, test meals, pharmacotherapies and medical devices, as well as dietary, exercise, and sleep interventions.

The appetite VAS typically include a series of 100-mm horizontal lines representing continua anchored at either end by the absence or extreme presence of a subjective appetite-related state [69]. Although the 100-mm VAS length is most commonly used, the 150-mm VAS is also valid and leads to similar results [70]. Therefore, appetite VAS scores can be obtained from 100-mm and 150-mm length scales both before and in response to a meal [70]. The participant indicates the point along each line that best represents how they feel at that moment.

The most commonly-used version of this questionnaire includes appetite dimensions specific to hunger, fullness, desire-to-eat, and prospective food consumption. Hunger and desire-to-eat are thought to be indices of a construct that reflect subjective hunger, while fullness and prospective food consumption are considered to reflect satiety. It is important that the differences between the four items are explained to participants prior to being asked to complete VAS ratings. A variant of the appetite VAS used in some European countries also adds a question on satiety [71]. Additionally, desire-to-eat can be expanded to desire-to-eat savoury/sweet foods, or other food types, as well as feelings of thirst, nausea and bloatedness. Description and specific questions pertaining to the five most commonly-used VAS are presented in Table 1.

To capture the immediate peak/nadir within meal responses, each appetite sensation should be assessed separately with measurements taken soon after arrival, immediately before and after the test meal, and at the following time-points: 10, 20, 30, 60, 90, 120, 150, and 180 min post-MM-TT [3]. Researchers may wish to perform additional assessments beyond this point, for example, if blood analytes are measured over a longer duration. Appetite VAS should be assessed before blood samples are obtained to reduce any influence of potential uncomfortable feelings during the insertion of a cannula or sample collection. Electronic devices can be accepted as a valid method for measuring appetite but should not be used interchangeably with paper and pen [72]. While data collected using the paper and pen method may not differ from that collected electronically, the former is more time-consuming as individual lines are measured manually and inputted into a spreadsheet, potentially introducing human error. When combined with the standardisation protocol outlined in this article, appetite sensations have been shown to be both valid and reproducible [23, 73]. The validity of appetite sensations as a predictor of energy intake has been demonstrated in numerous studies [74].

Presentation of appetite VAS scores

In addition to analysing single time-points or graphical curves, researchers have developed various formulae and statistical approaches to enhance the utility and sensitivity of the basic VAS scores. Participants' ratings for individual VAS (hunger, fullness, desire-to-eat, and prospective food consumption) are often analysed separately, but may be combined to summarise

Table 1. Description of perceived motivation to eat.

Dimension	Description	Question	Appetite state
Hunger	Captures immediate motivation to eat.	How hungry are you?	Ranging from "Not hungry at all" to "Extremely hungry"
Fullness	Reflects the feeling of being full or post-meal satiation.	How full are you?	Ranging from "Not full at all" to "Extremely full"
Desire-to-eat	Measures the general or specific DTE.	How strong is your desire to eat?	Ranging from "No desire to eat" to "Strong desire to eat"
Prospective food consumption	Measures the anticipated quantity of food that could be eaten.	What amount of food could you eat?	Ranging from "Nothing" to "A very large amount"
Satiety	Measures the feeling of being satiated or satisfied.	How satiated are you?	Ranging from "Not at all satiated" to "Extremely satiated"

DTE desire-to-eat, *PFC* prospective food consumption.

'motivation to eat' [75]. As such, when analysing VAS, it might be advisable to calculate a composite appetite score (CAS) based on all of the scale items, unless there is a specific hypothesis relating to individual items. Table 2 provides an in-depth summary of the most commonly presented VAS outcomes and their calculation, appropriate use, and example applications. Additional elaboration is given in the data handling section of this paper.

Additional considerations

Appetite VAS are cheap, easy to use, and simple to interpret. These scales demonstrate good within-subject reliability and validity [23] and broadly predict food intake [74]. Values also vary under conditions where they would be expected to change (e.g. in response to nutrient loads and anorectic drugs) [74]. However, the between-participant variability can be large, meaning VAS are more suitable for within-subject comparisons than between-subject comparisons. Robust familiarisation with VAS, and strict standardisation of procedures, before and during the MM-TT, limit the impact of confounding factors on VAS data. Importantly, VAS should not be used directly as a proxy for quantitative variables such as energy or food intake as they do not always predict these outcomes perfectly. As such, if energy or food intake is a primary outcome of the study, an *ad libitum* test meal should be used at the end of the MM-TT period.

The Three-Factor Eating Questionnaire (TFEQ) [76] and the Dutch Eating Behaviour Questionnaire (DEBQ) [77] are psychometric tools used in appetite studies to assess eating behaviours and attitudes towards food. As the constructs within these questionnaires (TFEQ: dietary restraint, disinhibition, hunger; DEBQ: emotional, external and restrained eating) can influence research outcomes, these tools are often employed during pre-screening to identify, screen and stratify individuals exhibiting extreme values. Additional eating behaviours or eating traits may also influence MM-TT outcomes, including but not limited to: food cravings, eating disorders, body image disturbance, mood/affective states, impulsivity/reward states, mindful eating and interoceptive awareness. Researchers should therefore use their judgement to determine whether other questionnaires should be employed before an MM-TT, to screen in/out participants or to retrospectively permit statistical adjustment and sensitivity analyses.

Although VAS are widely used in appetite research, it is important to acknowledge that they may be less suitable in certain situations. For example, difficulties may arise for individuals who have trouble understanding abstract concepts such as rating feelings on a scale, including those with visual impairments, cognitive impairments, low literacy levels, or language barriers. Young children (<7 years) are one prominent example, where more bespoke methods (e.g. picture-based scales) may be needed to assess appetite [78, 79].

BLOOD SAMPLING

Sampling blood before, and frequently after ingesting a MM-TT provides insight into metabolic and hormonal responses. This may enhance understanding of factors influencing appetite-related outcomes. The method of blood sampling (e.g. location, frequency and posture), can impact concentrations of circulating analytes and requires consideration.

Location of blood sampling

Location choice will depend on the research question and resources available. Generally, in physiological research, the most common locations of blood sampling are capillary beds (e.g. sampling from a fingertip, earlobe, or heel), and veins within the antecubital fossa of the forearm, providing mixed-venous blood. Capillary sampling is inexpensive and less technically challenging but limited by sample volume, making venous blood the common choice when several analytes are measured.

A more technically challenging method is sampling arterial blood, which provides the best indication of peripheral analyte availability, since arterial blood is delivered to most tissues. As tissue metabolism can alter the composition of the blood via net uptake, release or interconversion of an analyte, sampling non-arterial blood can misrepresent the peripheral availability of an analyte under some circumstances. The extent to which the sample location makes a difference to inferences drawn depends on factors such as the net tissue flux of the analyte and tissue blood flow. For example, after oral ingestion of a 75 g glucose load, there is a net uptake of glucose, insulin and GLP-1, and net release of lactate by the forearm, whereas the net forearm flux of triacylglycerol remains neutral (Fig. 2) [80–82]. Accordingly, for the most accurate representation of postprandial secretion and availability of analytes for the periphery, arterial or arterialised blood may be preferable to venous blood for many metabolites and hormones.

If arterial blood is most appropriate for the research question but is not practically feasible, then *arterialised* venous blood and capillary samples can replace arterial sampling for some analytes but not necessarily appetite related peptides. To obtain *arterialised* venous blood, retrograde cannulation of a pre-heated dorsal hand is needed and measuring oxygen saturation in blood samples can confirm arterialisation. The arterialisation method has been shown to provide samples representative of arterial blood under various conditions for some analytes; including, glucose, lactate, insulin, non-esterified fatty acids (NEFA), and beta-hydroxybutyrate [83–85]. However, it has been suggested that the method may affect some counter-regulatory hormone responses to hypoglycaemia (i.e., glucagon secretion and/or clearance) [86]. It is worth noting that in many circumstances, concentrations of analytes (e.g., glucose) within capillary samples are closer to arterial/arterialised than to venous samples. Therefore, capillary sampling can be an alternative to arterial/arterialised sampling if the required sample volume for analysis is relatively small. To our knowledge, representation has not been demonstrated in appetite-related peptides so further research is necessary to demonstrate if this is the case. Researchers are therefore encouraged to follow blood sampling guidelines in assay instructions meticulously.

Frequency and duration of blood sampling

The appropriate frequency and duration of blood sampling depends on the research question, resources, and study population, in addition to other aspects such as the size and composition of the MM-TT. Frequency of blood sampling should be sufficient to capture the required resolution of a postprandial response, ideally detecting a peak or nadir concentration of an analyte. Duration of sampling should ideally aim to capture the full postprandial response (with the return to fasting/basal concentrations) and at least capture a peak/nadir concentration.

In insulin-sensitive populations, glucose, insulin and NEFA concentrations typically achieve peak/nadir concentrations within 60 min and return to fasting concentrations within 120 min [87, 88]. In contrast, circulating ghrelin, gastric inhibitory polypeptide (GIP), GLP-1, and PYY concentrations may take longer than 180 min to return to basal [88–90]. The potential impact of sampling frequency on assessing a postprandial response is illustrated in Fig. 3, using data from published work [91]. An inference based on infrequent sampling would be that peak glucose concentrations would be achieved at 50 min after ingestion, and that prior breakfast consumption would result in no subsequent postprandial rise in glucose concentrations (Fig. 3A). Using the same data with increasing sampling frequency, an additional time-point of 30 min postprandially would result in an inference that peak glucose concentrations are achieved at 30 min in both conditions (Fig. 3B), and with 5-min sampling frequency, there is evidence of a difference in the time of peak glucose concentrations between

Table 2. Summary of outcomes, formulae and statistical approaches based on the appetite visual analogue scales (VAS).

Method	Description	Calculation	Reliability/validity	Discussion	Examples of application
Single time points	Measures appetite at specific moments such as fasting or pre-meal	Individual appetite VAS scores at distinct time points for each appetite sensation	Depends on timing and standardisation and does not reflect satiety response [3]. Often leads to many comparisons, increasing type 1 error.	Provides snapshots of appetite states for comparison. Could be used to compare pre-meal appetite differences in MM-TT.	Fasting studies, pre- and post-meal evaluations studies [23].
Multiple time points	Measures appetite over periods of time	Raw data or changes from baseline (baseline subtraction) for each appetite sensation	Method validated in many studies [3]. Baseline variables should be considered in the statistical analyses.	Changes from baseline are often preferred in MM-TT to account for pre-meal appetite differences and to isolate the postprandial response. However, both approaches can be used depending on the study objective and the nature of the intervention.	Used in studies to compare conditions [141].
Rate of recovery	Measures the time it takes for appetite levels to return to baseline after a meal or intervention	Change in appetite scores over time (e.g., slope of recovery curve) for each individual appetite sensation	Validity and reliability in specific contexts [142].	Useful in assessing the duration of satiety or hunger suppression	Used in food preloads, meal composition, and pharmacological agent studies [143, 144].
Early versus late satiety	Differentiates satiety responses at early versus later stages of the postprandial period	Comparison of appetite scores at early (mean of 0–60 min period) and late (mean of 60–180 min period) postprandial time points for each appetite sensation.	Method from Gibbons et al. [141], validated in satiety studies.	Highlights differences in satiety dynamics across time.	Used in meal structure and nutrient timing studies [141, 145]
Area Under the Curve (AUC)	Summarises the overall appetite response over a specified period	Calculated using trapezoidal method: the mean scores of pairs of adjacent time points are multiplied by the time interval between the sampling points [125]. Total AUC includes the total area between the origin line (zero) and response curve; incremental AUC is restricted to areas above a nominal threshold (most often baseline) [126, 128, 146]	Trapezoidal method is widely used and validated [126, 146].	Provides a comprehensive measure of appetite over time representative of the magnitude and duration of response. Incremental AUC may be more appropriate if baseline differences are evident and/or the response to the meal is the primary interest.	Applied in preload and intervention studies [72].
Composite Appetite Score (CAS)	Combines multiple appetite measures (e.g., hunger, fullness, desire-to-eat, prospective food consumption) into a single index	Average or weighted combination of individual appetite scores (usually four dimension – ([hunger + DTE + PFC + (100−fullness)]/4)	Method validated in multiple contexts [147].	Enhances sensitivity and utility of appetite data.	Studies requiring a global appetite measure [148, 149].
Satiety Quotient (SQ)	Evaluates the satiety efficiency in response to food intake.	Change in appetite scores divided by energy content of the meal [Mean 1 h post appetite scores-pre/kcal MM-TT*100]. Can be used for an individual appetite sensation or average appetite scores (e.g. hunger, fullness, DTE and PFC).	Method from Green et al. [150], validated against reported and measured energy intakes [151].	Links subjective appetite to objective energy intake, emphasising food quality and quantity.	Used to document the strength of satiety related to biology of the individual satiety efficiency research, preload studies, and portion control interventions [152, 153].

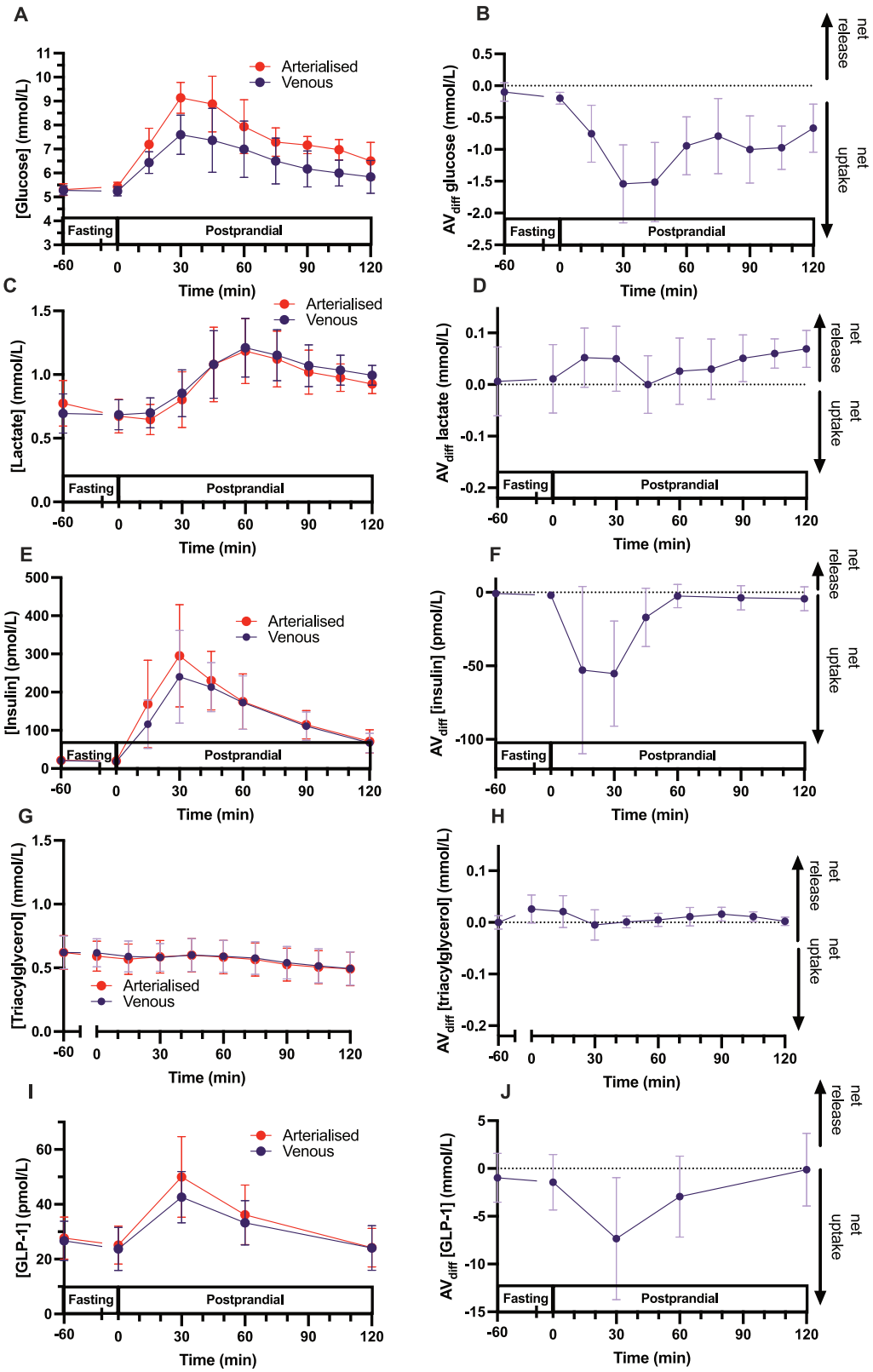


Fig. 2 Metabolite responses to a standard oral glucose tolerance test at different sampling locations. Plasma glucose (**A**, **B**), lactate (**C**, **D**), insulin (**E**, **F**), triacylglycerol (**G**, **H**), and glucagon-like peptide 1 (GLP-1; **I**, **J**) concentrations in arterialised and mixed venous blood (**A**, **C**, **E**, **G**, **I**) and net forearm uptake/release (**B**, **D**, **F**, **H**, **J**) before and after ingestion of 75 g glucose. Data redrawn from https://doi.org/10.1017/S0007114517001362 and https://doi.org/10.1113/EP087118.

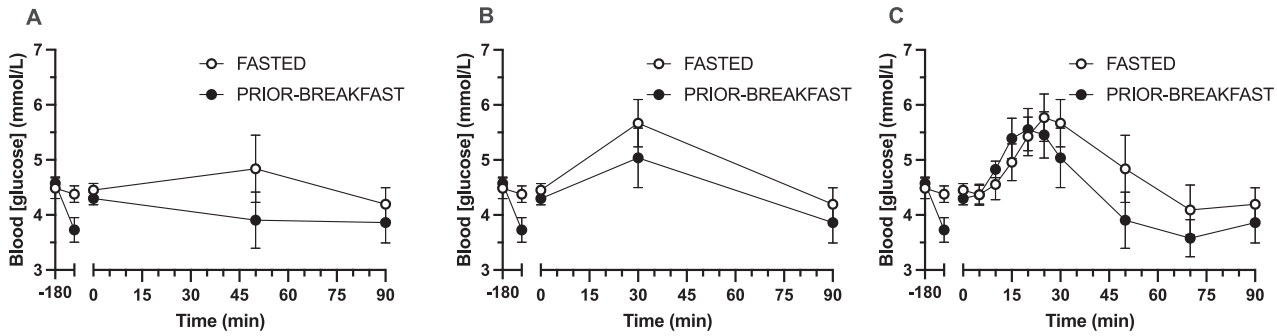


Fig. 3 Circulating blood glucose profiles in response to fasting and breakfast consumption at different sampling intervals. Panels show circulating blood glucose concentrations with blood sampling conducted at 50 min (**A**), 30 min (**B**) and at 5-min intervals (**C**). Data taken from: https://doi.org/10.1017/S0007114512005582.

conditions (20 vs. 25 min; Fig. 3C). A summary of recommendations regarding sampling location, frequency and duration can be found in Table 3, although researchers may have valid reasons for alternatives to these recommendations.

Posture

Changing the volume of distribution may affect the measured concentration even if the absolute amount of the analyte has not changed. Changing posture (e.g., supine versus sitting versus standing), impacts plasma volume and thus the concentrations measured. For example, changing posture from standing to supine increases plasma volume by >400 mL within 30 min, and return to standing decreases plasma volume by >500 mL within 30 min [92]. It is therefore recommended that posture is controlled (with a stabilisation phase of at least 30 min before sampling), or that plasma volume is measured to correct for changes in the volume of distribution.

APPETITE-RELATED PEPTIDES

Appetite-related peptides communicate acute and chronic energetic/nutritional status to the central nervous system (Fig. 4). These neurohumoral messengers (peptides) include leptin—a tonic signal—which informs the brain about energy balance and available energy reserves. Concomitantly, ghrelin, cholecystokinin (CCK), GIP, GLP-1 and PYY primarily function in an episodic nature; influencing appetite sensations, eating behaviour and digestive processes in response to individual meals.

Certain appetite-related peptides are prone to degradation (in vivo and ex vivo) and circulate as different isoforms with divergent effects. Researchers must therefore consider which peptides to measure and ensure that blood sampling procedures maintain peptide integrity. Generic recommendations are listed in Table 4. Bespoke considerations for measuring episodic peptides are discussed herein. Tonic adiposity-related peptides, such as leptin and insulin, are less challenging to measure and are not addressed here. Likewise, glucagon [93] and CCK [94] measurements have been reviewed in detail elsewhere.

Ghrelin

Ghrelin is a 28 amino acid peptide most prominently secreted from the stomach. It is unique in being O-acylated at its third residue (serine) [95]. Ghrelin O-acyltransferase catalyses post-translational ghrelin acylation, permitting receptor binding and appetite effects [96]. Therefore, ghrelin circulates as acylated (AG) and unacylated (UAG) peptides. As UAG has diverging effects to AG, measurement of AG is recommended in appetite research. During blood sampling, this necessitates preservation of the octanoyl moiety which is prone to protease degradation [97].

Studies have investigated how sample processing affects circulating AG [18, 98–101]. Recommendations include the necessity to keep samples chilled; with pre-cooling of blood collection tubes and refrigerated centrifugation [101]. Samples should be centrifuged immediately or placed on ice for no longer than 15 min [99]. Protease inhibitor(s) must be added to samples; ideally to blood collection tubes to ensure immediate mixing [98, 99, 101]. Protease inhibitor choice should be based on assay recommendations. Sample acidification (hydrochloric acid) may be recommended for some assays to enhance preservation with long-term storage and freeze-thaw cycles [100].

Incretins: GIP and GLP-1

GIP is a 42 amino acid peptide primarily secreted by the small intestine in its intact isoform, $GIP_{1\text{-}42}$ [102]. Upon secretion, it undergoes rapid degradation (~5–7 min half-life) by dipeptidyl peptidase-4 (DPP-4) to form the inactive metabolite $GIP_{3\text{-}42}$ [103], the predominant circulating isoform. In applied studies where GIP secretory responses to different interventions are of interest, it is recommended that researchers measure 'total GIP' [104]. However, GIP may act solely in an endocrine fashion, therefore researchers must consider whether the intact isoform alone may be of interest [104].

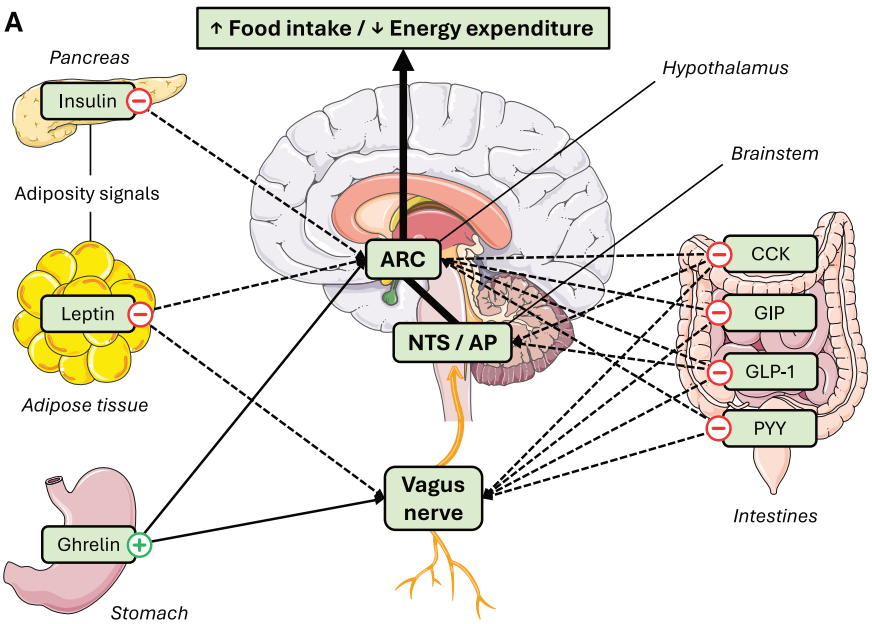
GLP-1 is primarily secreted from the intestine as two 'intact' isoforms: glycine-extended $GLP\text{-}1_{7\text{-}37}$ and amidated $GLP\text{-}1_{7\text{-}36}$ amide [102]. Like GIP, degradation by DPP-4 occurs (~1–2 min half-life) to form the primary metabolites, $GLP\text{-}1_{9\text{-}37}$ and $GLP\text{-}1_{9\text{-}36}$ amide [105, 106]. In humans, the amidated isoforms predominate [105, 107]. Only 10% of secreted GLP-1 reaches the circulation intact [106]; therefore, unlike GIP, intact GLP-1's effects on satiety are predominantly through vagal afferents before degradation and systemic release [108]. Whilst the GLP-1 metabolites were initially thought to be inactive, this notion has been contested [109]. Given that the human intestine secretes 100% of GLP-1 in its intact form [110], and its rapid degradation alters concentrations between the secretion and measurement site (when sampling venous blood), peripheral concentrations of intact GLP-1 may not rise postprandially despite significant secretion [111]. It is similarly recommended to measure total GLP-1 (i.e. both $GLP\text{-}1_{7\text{-}36}$ amide and $GLP\text{-}1_{9\text{-}36}$ amide) when estimating GLP-1 secretory responses [104]. Conversely, measurement of intact GLP-1 alone may be relevant to specific scenarios, such as assessing responses to different DPP-4 inhibitors [104].

When measuring intact GIP and GLP-1, addition of a DPP-4 inhibitor to collection tubes before sample acquisition helps prevent ex vivo degradation [112]. Furthermore, ethanol or solid-phase extraction procedures can reduce non-specific interference by other plasma proteins [113]. These result in lower circulating concentrations and accentuate postprandial responses, enabling greater reproducibility across individuals, assays and studies [113]. These procedures are not necessary when measuring total GIP and GLP-1 [104]. Traditionally, aprotinin and blood collection tubes containing protease inhibitor cocktails (e.g. BD™ P800) have been

Table 3. Recommended priority sampling location, frequency and duration for various analytes.

Analyte	Recommended sampling location	Ideal sampling frequency (min)	Ideal sampling duration (min)
CCK [145]	Arterialised, Capillary, or Mixed Venous.[a]	≤15 min	≥180 min
Glucagon [154]	Arterialised, Capillary, or Mixed Venous.[a]	≤15 min	≥240 min
Ghrelin	Arterialised or Capillary.[a]	≤15 min	≥240 min
GIP [155]	Arterialised, Capillary, or Mixed Venous.[a]	≤15 min	≥240 min
GLP-1 [155]	Arterialised or Capillary.	≤15 min	≥240 min
Leptin[b] [155]	Arterialised, Capillary, or Mixed Venous.[a]	≤240 min	≥360 min
PYY [156]	Arterialised or Capillary.[a]	≤15 min	≥360 min

CCK cholecystokinin, *GIP* glucose-dependent insulinotropic polypeptide, *GLP-1* glucagon-like peptide-1, *PYY* peptide-YY.
[a]Lack of evidence for a clear recommendation.
[b]Leptin often only measured in the fasted state.



B

Insulin	Leptin	Ghrelin	CCK	GIP	GLP-1	PYY
↓ food intake from chronic adiposity signals	↓ food intake from chronic adiposity signals	↑ food intake & appetite	↓ food intake & appetite	↓ food intake & appetite?	↓ food intake & appetite	↓ food intake & appetite
↓ circulating glucose	↑ satiation signals from CCK	↑ gastric acid secretion & emptying	↓ gastric acid secretion & emptying	↓ gastric acid secretion & emptying	↓ gastric acid secretion & emptying	↓ gastric acid secretion & emptying
↑ glycogen synthesis	↓ insulin secretion	↓ insulin secretion	↑ pancreatic enzyme secretion	↑ glucose-dependent insulin secretion*	↑ glucose-dependent insulin secretion	↑ insulin secretion
↓ lipolysis & proteolysis	↓ glucagon secretion	↑ adiposity	↑ gallbladder contraction	↑ glucagon secretion*	↓ glucagon secretion	
↑ lipogenesis	↑ energy expenditure	↑ pituitary release of GH		↑ lipogenesis and lipid buffering capacity	↑ pancreatic β-cell proliferation	

Fig. 4 Overview of the role of appetite-related peptides in the neurohumoral control of food intake and energy expenditure. A Central and peripheral pathways involved in the appetite-related peptide control of food intake and energy expenditure. **B** Physiological functions of different appetite-related peptides. (+, solid line) indicates a positive effect on increasing food intake and/or decreasing energy expenditure; (−,dashed line) indicates a negative effect. * Indicates these effects may be blunted in conditions of hyperglycaemia. AP area postrema, ARC arcuate nucleus, CCK cholecystokinin, GH growth hormone, GIP gastric inhibitory polypeptide, GLP-1 glucagon-like peptide-1, NTS nucleus tractus solitarius, PYY peptide YY.

Table 4. Generic blood sampling and processing considerations prior to measurement of appetite-related peptides.

Recommendation	Explanation
Standardise participants' body position for all blood draws (at least 20 min before each collection)	Body posture influences blood and plasma volume. The concentration of analytes in circulation may therefore vary if blood is drawn in different body positions (e.g. seated versus supine).
Check assay manufacturer guidelines when choosing blood sample collection tubes (e.g. EDTA, lithium heparin, clotting activator).	Peptide integrity is better preserved with some sample preservatives than others. Some preservatives interfere with assay biochemistry.
Pre-cool blood sample collection tubes by placing in wet ice prior to blood draws.	Ex vivo sample degradation is reduced at lower temperatures.
Use best practice phlebotomy techniques to avoid blood sample haemolysis (i.e. avoid excessively tight tourniquet and release immediately as blood flows, handle samples carefully, prompt centrifugation).	Haemolysis of blood samples may interfere in assay biochemistry and lead to missing data.
Centrifuge samples immediately after collection. If not possible, place samples in wet ice and process within 15 min.	Ex vivo sample degradation is reduced once plasma/serum is harvested.
Prepare single use sample aliquots to avoid assaying peptides after freeze-thaw cycles.	The freeze-thaw process may damage protein structure (particularly repeated freeze-thaws).
For long-term storage, freeze sample aliquots at −80 °C.	Colder sample storage (i.e. versus −20 °C) is likely to better preserve protein integrity.
Analyse samples in duplicate.	Duplicate analyses help to enhance result precision.
Analyse all samples from the same participant on the same assay plate/run. Within a research study, ensure all kits are from the same assay batch number where possible.	Inter-assay co-efficient of variation is greater than intra-assay co-efficient of variation.
Check the dynamic/working range when choosing appetite-related hormone assays.	Absolute concentrations of peptides may vary considerably depending on factors such as participant feeding status (fed versus fasted) and adiposity (obesity versus healthy weight).

used to further prevent degradation; however, recent investigations concluded EDTA-coated tubes without protease inhibitors are sufficient for clinical trials [114]. Immediate sample processing may obviate the need for these protease inhibitors.

Commercially available GLP-1 enzyme-linked immunosorbent assays (ELISA) were recently compared to an established in-house RIA [115]. Considerable variation in assay performance was observed between different assays and batches of the same assay. Therefore, analysing samples within the same batch using a highly specific and sensitive assay for the isoform(s) of interest is important [107]. Whilst radio-immunoassays were traditionally used, adequate ELISAs for intact and total GIP are commercially available [114]. Nevertheless, researchers should note that a short-form of GIP has recently been identified ($GIP_{1\text{-}30}$ amide) which may possess similar biological activity to full-length GIP but is not detected by many current assays [112]. This could potentially underestimate GIP responses, however, specific assays for this isoform are under development [112].

PYY

PYY is a 36 amino acid peptide secreted most prominently from L-cells in the distal intestine. In circulation, PYY exists in two major forms, as full-length $PYY_{1\text{-}36}$ and as a 34 amino acid peptide ($PYY_{3\text{-}36}$) [116]. The truncated peptide is produced by enzymatic cleavage of two amino-terminal peptides, by DPP-4 [117]. This step is crucial as only $PYY_{3\text{-}36}$ can bind to the neuropeptide-Y (NPY) Y2-receptor, mediating appetite [97]. Sample collection and processing procedures must therefore inhibit the ex vivo conversion of $PYY_{1\text{-}36}$ to $PYY_{3\text{-}36}$ and further protease degradation.

Assay manufacturers typically recommend that samples destined for $PYY_{3\text{-}36}$ analysis are treated with DPP-4 inhibitor and aprotinin (broad spectrum protease inhibitor). In experiments, blood is drawn into syringes containing DPP-4 inhibitor to block peptide truncation immediately, with aprotinin subsequently added. However, one study found that neither DPP4 inhibitor or aprotinin influenced fasted, fed, or postprandial concentrations of $PYY_{3\text{-}36}$ [18]. The immediacy of sample processing (centrifugation) may explain this outcome and it is unclear how repeated freeze/thaw cycles would have impacted findings. EDTA and aprotinin treatment do, however, inhibit the degradation of $PYY_{3\text{-}36}$ to $PYY_{3\text{-}34}$ (inactive metabolite) [118]. If $PYY_{3\text{-}36}$ cannot be measured, $PYY_{1\text{-}36}$ and $PYY_{3\text{-}36}$ broadly track one another from the fasted to postprandial state, meaning circulating levels of total PYY provide a reasonable marker of overall PYY secretion. However, there are scenarios where the ratio of $PYY_{1\text{-}36}$ and $PYY_{3\text{-}36}$ is altered [119]; therefore, assays directed at $PYY_{3\text{-}36}$ should be prioritised when accurate quantification of the appetite-related isoform is the aim.

DATA HANDLING

Data presentation

Data from the MM-TT are characterised by the appetite-related outcome(s) quantified before and at multiple time-points after the meal stimulus. Appetite data are usually collected from two or more MM-TTs, showing data from participants studied under different conditions, different points in time and/or different populations.

Variables with multiple postprandial assessments are often presented graphically as a response curve over time (Fig. 5A). As an alternative to raw values, appetite perceptions and appetite-related hormone concentrations can be presented and analysed as changes from baseline (Fig. 5B and Table 2). This may minimise the influence of day-to-day variability and improve sensitivity to examine predictors of interest [120]. For outcomes limited to single or possibly duplicate assessments, or those converted to summary values (see below), raw data may be better suited to presentation in table format or in visual plots that display the data distribution [121, 122] (Fig. 5C, D).

Data analysis

Raw data from every sampling time-point during the MM-TT can be integrated into a statistical model as an independent variable with other predictors such as condition or group. This allows examination of the mean effect of each independent variable on the appetite outcome, averaging across the levels of the other independent variables. The main effect of time can identify

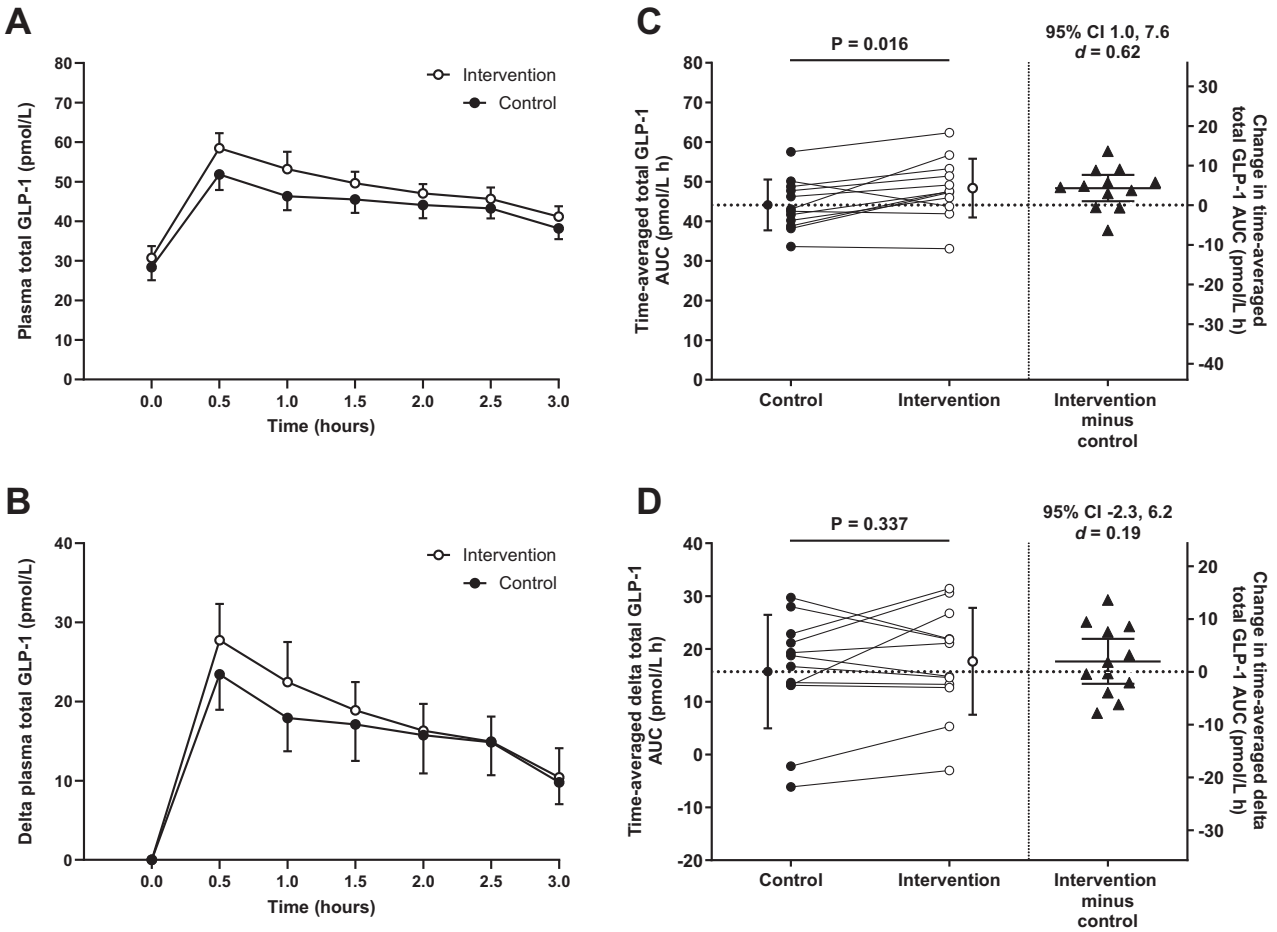


Fig. 5 Plasma glucagon-like peptide-1 (GLP-1) concentrations during a 3-h mixed meal tolerance test with the meal provided after the fasted sample at 0-h. Data are hypothetical for $n = 12$ in a randomised crossover study involving an intervention and control study arm. On left hand panels, data presented as **A** raw analysed concentrations or **B** change from baseline (delta) and displayed as mean (SEM) to prevent figure distortion. On right hand panels **C**, **D**, the left hand *y* axis displays the mean (SD) time-averaged total area under the curve (AUC) and circles with connecting lines show the individual participant values. The right hand *y* axis displays the mean difference (95% confidence interval) between the intervention and control arms and the triangles represent the individual participant data. The *p* value, 95% confidence interval and standardised effect size (Cohen's *d*) is presented for the mean difference between the intervention and control study arms.

differences in the time course of responses across the level of another independent variable when integrated into an interaction term.

Researchers are often less concerned with the influence of sampling time so it may be preferable to remove the time series element by calculating a summary statistic. This approach allows for direct comparisons of summary condition or group means, with adjustment for baseline (pre-intervention) as a covariate recommended for longitudinal study designs [123, 124]. Temporal data can be condensed into a summary statistic by calculating the area under the curve (AUC) using the trapezoidal rule. The total AUC involves summing the AUC calculated between the origin line at zero and the response curve for each consecutive sampling point, generating a single value representing a time-weighted average of the postprandial response [125]. This can be converted to a time-averaged value by dividing by the observation period which is often easier to interpret (Fig. 5C).

Whilst the total AUC indicates the overall response, it does not adjust for baseline variability (unless calculated using delta values [Fig. 5D]). Alternatively, the incremental AUC can be quantified using the same method as the total AUC but is restricted to areas above a nominal threshold whilst ignoring areas below the threshold [126, 127]. The threshold adopted is commonly the fasting value, but for outcomes such as acylated ghrelin that decrease in response to meals, the nadir value may be more appropriate. For the MM-TT, appetite responses to the meal are the primary focus and since baseline (fasting) values can vary, incremental AUC may be preferred over total AUC to better isolate meal-induced effects [128]. The freely available *Time Series Response Analyser* provides an automated tool for the accurate calculation of several summary statistics [128]. In addition to AUC, peak or nadir and time-to-peak or time-to-nadir values can be calculated for each response curve providing sufficient sampling time-points are collected to generate precise estimates.

Data considerations—sample size and target difference

No clinically relevant target differences have been established for appetite parameters. Establishing an important difference is central to sample size estimations and interpreting study findings and can be guided by relevant literature [129]. For example, a 10-mm (10%) difference in mean or AUC appetite has been proposed as a realistic difference between two meals [23]. Difference estimates can be supplemented by standardised effect sizes often expressed for continuous outcomes using Cohen's *d* thresholds (0.2, 0.5 and 0.8 for small, medium, and large effects, respectively) [130]. Clinical or practical relevance of MM-TT findings should not

depend solely on statistical significance but also be informed by the target difference and precision of treatment estimates reflected in confidence intervals [129, 131].

Data considerations—missing data

Missing MM-TT data is almost inevitable when multiple time-sensitive measures are collected on repeat occasions. Comprehensive guidelines exist for handling missing data in randomised controlled trials [132–134] and many of these principles hold for time series data collected in MM-TTs. Single imputation methods using the sample mean or regression-based predictions may be appropriate when missingness is completely random (e.g., a missing blood sample due to sampling issues). A tool has been developed to simulate missing MM-TT data that accounts for the correlated nature of time series data and could be useful when ≤2 consecutive values are missing [135]. Other approaches may be required if missing data can be explained by observed variables (e.g., greater missing samples in women than men) or if missingness is systematically dependent on unobserved data (e.g., less follow-up data in those benefiting less from an intervention) [132]. These include multiple imputation and mixed effects models which can be supplemented by sensitivity analyses to assess the impact of missing data on treatment estimates [132].

Data considerations—outliers

Outliers can be identified visually from boxplots, histograms, or scatterplots, and/or mathematically using the mean or median. For example, standardised residuals, calculated by dividing the difference between the observed and mean predicted value by its standard deviation, beyond a threshold of 3.5 has been proposed for sample sizes <50, >4 for samples up to 500, >4.5 for samples up to 5000, and so forth [131]. Alternatively, a threshold of the median ±2.5 times the median deviation of scores from the median is likely to be less sensitive to the influence of outliers than calculation of the mean and standard deviation [136]. Outliers should not be excluded automatically without strong evidence that the outlier is not representative of the population of interest. Findings should be presented with and without the outlier(s) included and with full transparency on the process adopted for outlier identification [137].

Inter-individual variability

Although most studies quantify mean treatment or group effects, inter-individual variability may exist in appetite responses to identical meals beyond within-subject random variation [138]. Inter-individual variability in the magnitude of postprandial appetite suppression is an important consideration but requires robust methods to quantify. This can be established by replication of treatment and control arms (replicate crossover design) allowing separation of individual response heterogeneity from measurement error and/or natural fluctuation in outcomes over time [139, 140].

CONCLUDING REMARKS AND DEVELOPMENT OF THE MM-TT CHECKLIST

This paper has sought to highlight the plethora of factors that require consideration when designing MM-TTs. We hope this paper can be used as a resource by researchers working in relevant fields, helping to support decisions about MM-TT design in the planning stages of research. While the primary focus of this paper is the appetite research field, various elements are relevant to those using MM-TTs in other research contexts. By scrutinising all elements of the MM-TT, this paper should help researchers to better understand the MM-TT process including the ability to recognise strengths and limitations in its application. To support these ambitions, we have developed a checklist (Supplementary Table 1) which pools key factors to consider when designing a MM-TT. Researchers can use this list to evaluate their choice of methods and we encourage its adaptation for use in wider fields of research.

It is important to note that this paper did not seek to identify a 'gold standard' MM-TT to be recommended for all. This is because the MM-TT technique is used in a range of contexts, with varied research aims and available resources. Instead, our review has highlighted key factors which underpin the adaptability of this method. Also relevant to note, in some cases we have been unable to make clear recommendations given lack of evidence. There are many examples of this regarding familiarisation and standardisation procedures, for instance. Moreover, for many issues the methodological approach depends on the research question. Finally, the scope of this paper was limited to core MM-TT factors. Further discussion about wider, related constructs and techniques (e.g. measurement of energy intake, diet-induced thermogenesis, gastric emptying) is needed in future work.

DATA AVAILABILITY

This review paper does not contain any original data. All information presented is derived from previously published studies, which are cited appropriately within the text. As such, no new data were generated or analysed during the preparation of this manuscript.

REFERENCES

1. Casanova N, Finlayson G, Blundell JE, Hopkins M. Biopsychology of human appetite—understanding the excitatory and inhibitory mechanisms of homeostatic control. Curr Opin Physiol. 2019;12:33–8.
2. Beaulieu K, Blundell J. The psychobiology of hunger—a scientific perspective. Topoi. 2021;40:565–74.
3. Blundell J, De Graaf C, Hulshof T, Jebb S, Livingstone B, Lluch A, et al. Appetite control: methodological aspects of the evaluation of foods. Obes Rev. 2010;11:251–70.
4. Gibbons C, Hopkins M, Beaulieu K, Oustric P, Blundell JE. Issues in measuring and interpreting human appetite (satiety/satiation) and its contribution to obesity. Curr Obes Rep. 2019;8:77–87.
5. Oustric P, Thivel D, Dalton M, Beaulieu K, Gibbons C, Hopkins M, et al. Measuring food preference and reward: application and cross-cultural adaptation of the Leeds Food Preference Questionnaire in human experimental research. Food Qual Prefer. 2020;80:2019.103824.
6. Blundell J, Rogers P, Hill A. Evaluating the satiating power of foods: implications for acceptance and consumption. In: Colms J, Booth DA, Pangborn RM, Raunhardt O. (eds) Food acceptance and nutrition. London: Academic Press; 1987. pp. 205–19.
7. Mittelman SD, Klier K, Braun S, Azen C, Geffner ME, Buchanan TA. Obese adolescents show impaired meal responses of the appetite-regulating hormones ghrelin and PYY. Obesity. 2010;18:918–25.
8. Clamp L, Hehir APJ, Lambert EV, Beglinger C, Goedecke JH. Lean and obese dietary phenotypes: differences in energy and substrate metabolism and appetite. Br J Nutr. 2015;114:1724–33.
9. Beaulieu K, Hopkins M, Blundell J, Finlayson G. Does habitual physical activity increase the sensitivity of the appetite control system? A systematic review. Sports Med. 2016;46:1897–919.
10. Yanni AE, Konstantopoulos P, Kartsioti K, Binou P, Karathanos V, Chatzigeorgiou A, et al. Effects of 12-week, non-energy-restricted dietary intervention with conventional yogurt on appetite hormone responses of type 2 diabetic patients. Food Sci Nutr. 2021;9:6610–6.
11. Martins C, Kulseng B, King NA, Holst JJ, Blundell JE. The effects of exercise-induced weight loss on appetite-related peptides and motivation to eat. J Clin Endocrinol Metab. 2010;95:1609–16.
12. van den Broek M, de Heide LJM, Emous M, Wijma RB, Veeger NJGM, Wolthuis A, et al. Satiety and gastrointestinal hormones during a Mixed Meal Tolerance Test after gastric bypass surgery: association with plasma amino acid concentrations. Surg Obes Relat Dis. 2018;14:1106–17.
13. Sargeant JA, King JA, Yates T, Redman EL, Bodicoat DH, Chatterjee S, et al. The effects of empagliflozin, dietary energy restriction, or both on appetite-regulatory gut peptides in individuals with type 2 diabetes and overweight or obesity: the SEESAW randomized, double-blind, placebo-controlled trial. Diab Obes Metab. 2022;24:1509–21.
14. Martin C, Ravussin E, Sanchez-Delgado G, Nishiyama H, Li J, Urva S, et al. The effect of tirzepatide during weight loss on food intake, appetite, food preference and food craving in people with obesity. Can J Cardiol. 2024;39:S239–S239.
15. Blundell J, Finlayson G, Axelsen M, Flint A, Gibbons C, Kvist T, et al. Effects of once-weekly semaglutide on appetite, energy intake, control of eating, food preference and body weight in subjects with obesity. Diab Obes Metab. 2017;19:1242–51.

16. Koopman ADM, Rutters F, Rauh SP, Nijpels G, Holst JJ, Beulens JW, et al. Incretin responses to oral glucose and mixed meal tests and changes in fasting glucose levels during 7 years of follow-up: the Hoorn Meal Study. PLoS ONE. 2018;13:0191114.
17. Skantze V, Hjorth T, Wallman M, Brunius C, Dicksved J, Pelve EA, et al. Differential responders to a mixed meal tolerance test associated with type 2 diabetes risk factors and gut microbiota—data from the MEDGI-Carb randomized controlled trial. Nutrients. 2023;15:4369.
18. Chandarana K, Drew ME, Emmanuel J, Karra E, Gelegen C, Chan P, et al. Subject standardization, acclimatization, and sample processing affect gut hormone levels and appetite in humans. Gastroenterology. 2009;136:2115–26.
19. Adam TC, Epel ES. Stress, eating and the reward system. Physiol Behav. 2007;91:449–58.
20. Scheer FAJL, Morris CJ, Shea SA. The internal circadian clock increases hunger and appetite in the evening independent of food intake and other behaviors. Obesity. 2013;21:421.
21. James LJ, Maher T, Biddle J, Broom DR. Eating with a smaller spoon decreases bite size, eating rate and *ad libitum* food intake in healthy young males. Br J Nutr. 2018;120:830–7.
22. Wasse LK, King JA, Stensel DJ, Sunderland C. Effect of ambient temperature during acute aerobic exercise on short-term appetite, energy intake, and plasma acylated ghrelin in recreationally active males. Appl Physiol Nutr Metab. 2013;38:905–9.
23. Flint A, Raben A, Blundell JE, Astrup A. Reproducibility, power and validity of visual analogue scales in assessment of appetite sensations in single test meal studies. Int J Obes. 2000;24:38–48.
24. Corney RA, Sunderland C, James LJ. The effect of hydration status on appetite and energy intake. J Sports Sci. 2015;33:761–8.
25. Corney RA, Sunderland C, James LJ. Immediate pre-meal water ingestion decreases voluntary food intake in lean young males. Eur J Nutr. 2016;55:815–9.
26. Yeomans MR. Short term effects of alcohol on appetite in humans. Effects of context and restrained eating. Appetite. 2010;55:565–73.
27. Schutz Y. Role of substrate utilization and thermogenesis on body-weight control with particular reference to alcohol. Proc Nutr Soc. 2000;59:511–7.
28. Dorling J, Broom DR, Burns SF, Clayton DJ, Deighton K, James LJ, et al. Acute and chronic effects of exercise on appetite, energy intake, and appetite-related hormones: the modulating effect of adiposity, sex, and habitual physical activity. Nutrients. 2018;10:E1140.
29. Stocks B, Zierath JR. Post-translational modifications: the signals at the intersection of exercise, glucose uptake, and insulin sensitivity. Endocr Rev. 2022;43:654–77.
30. Schmid SM, Hallschmid M, Jauch-Chara K, Born J, Schultes B. A single night of sleep deprivation increases ghrelin levels and feelings of hunger in normal-weight healthy men. J Sleep Res. 2008;17:331–4.
31. Magee CA, Huang X-F, Iverson DC, Caputi P. Acute sleep restriction alters neuroendocrine hormones and appetite in healthy male adults. Sleep Biol Rhythms. 2009;7:125–7.
32. Spiegel K, Tasali E, Penev P, Cauter EV. Brief Communication: Sleep curtailment in healthy young men is associated with decreased leptin levels, elevated ghrelin levels, and increased hunger and appetite. Ann Intern Med. 2004;141:846.
33. McNeil J, Forest G, Hintze LJ, Brunet J-F, Finlayson G, Blundell JE, et al. The effects of partial sleep restriction and altered sleep timing on appetite and food reward. Appetite. 2017;109:48–56.
34. Chao AM, Wadden TA, Ashare RL, Loughead J, Schmidt HD. Tobacco smoking, eating behaviors, and body weight: a review. Curr Addict Rep. 2019;6:191–9.
35. Kechter A, Ceasar RC, Simpson KA, Schiff SJ, Dunton GF, Bluthenthal RN, et al. A chocolate cake or a chocolate vape? Young adults describe their relationship with food and weight in the context of nicotine vaping. Appetite. 2022;175:106075.
36. El.Golli N, Dkhili H, Dallagi Y, Rahali D, Lasram M, Bini-Dhouib I, et al. Comparison between electronic cigarette refill liquid and nicotine on metabolic parameters in rats. Life Sci. 2016;146:131–8.
37. Breit M, Hoskinson H, Pitzer C, Wu Z, Bryner R, Olfert IM. Effects of electronic cigarette vapor on body mass, food intake, and body composition. FASEB J. 2017;31:1037.6.
38. Foltin RW, Fischman MW, Byrne MF. Effects of smoked marijuana on food intake and body weight of humans living in a residential laboratory. Appetite. 1988;11:1–14.
39. Anthony JC, Warner LA, Kessler RC. Comparative epidemiology of dependence on tobacco, alcohol, controlled substances, and inhalants: basic findings from the National Comorbidity Survey. Exp Clin Psychopharmacol. 1994;2:244–68.
40. Grotenhermen F. Pharmacokinetics and pharmacodynamics of cannabinoids. Clin Pharmacokinet. 2003;42:327–60.
41. Sexton M, Cuttler C, Mischley LK. A survey of cannabis acute effects and withdrawal symptoms: differential responses across user types and age. J Altern Complement Med. 2019;25:326–35.
42. Harpaz E, Tamir S, Weinstein A, Weinstein Y. The effect of caffeine on energy balance. J Basic Clin Physiol Pharm. 2017;28:1–10.
43. Patwardhan R, Demond P, Johnson R, Schenker S. Impaired elimination of caffeine by oral contraceptive steroids. J Lab Clin Med. 1980;95:603–8.
44. Tucker JAL, McCarthy SF, Bornath DPD, Khoja JS, Hazell TJ. The effect of the menstrual cycle on energy intake: a systematic review and meta-analysis. Nutr Rev. 2025;83:e866–e876.
45. Rogan MM, Black KE. Dietary energy intake across the menstrual cycle: a narrative review. Nutr Rev. 2023;81:869–86.
46. Geary N, Asarian L, Leeners B. Best practices for including sex as a variable in appetite research. Appetite. 2025;207:107840.
47. Schmalenberger KM, Tauseef HA, Barone JC, Owens SA, Lieberman L, Jarczok MN, et al. How to study the menstrual cycle: practical tools and recommendations. Psychoneuroendocrinology. 2021;123:104895.
48. Metz L, Isacco L, Redman LM. Effect of oral contraceptives on energy balance in women: a review of current knowledge and potential cellular mechanisms. Metabolism. 2022;126:154919.
49. Gill H, Gill B, El-Halabi S, Chen-Li D, Lipsitz O, Rosenblat JD, et al. Antidepressant medications and weight change: a narrative review. Obesity. 2020;28:2064–72.
50. Johannsen DL, Galgani JE, Johannsen NM, Zhang Z, Covington JD, Ravussin E. Effect of short-term thyroxine administration on energy metabolism and mitochondrial efficiency in humans. PLoS ONE. 2012;7:e40837.
51. Mohseni M, van der Valk ES, Van der Hurk MJB, Savas M, Boon MR, van Rossum EFC. Corticosteroid use and long-term changes in weight and waist circumference: the Lifelines Cohort study. J Clin Endocrinol Metab. 2025. dgaf166.
52. Melson E, Ashraf U, Papamargaritis D, Davies MJ. What is the pipeline for future medications for obesity?. Int J Obes. 2024;49:433–51.
53. Apovian CM, Okemah J, O'Neil PM. Body weight considerations in the management of type 2 diabetes. Adv Ther. 2019;36:44–58.
54. Fenton JM, King JA, McLeod CJ, Hoekstra SP, Finlayson G, Goosey-Tolfrey VL. A comparison of meal-related appetite, food reward and eating behaviour traits in people with and without spinal cord injury. Appetite. 2023;181:106384.
55. Siddall AG, Powell SD, Needham-Beck SC, Edwards VC, Thompson JES, Kefyalew SS, et al. Validity of energy expenditure estimation methods during 10 days of military training. Scand J Med Sci Sports. 2019;29:1313–21.
56. de Lima Macena M, da Costa Paula DT, da Silva AE, Praxedes DRS, Pfrimer K, Ferriolli E, et al. Use of the activPAL® triaxial accelerometer to estimate total energy expenditure in low-income women: differences between body mass index classifications. Arch Endocrinol Metab. 2023;67:e000616.
57. Westerterp KR. Physical activity and physical activity induced energy expenditure in humans: measurement, determinants, and effects. Front Physiol. 2013;4:90.
58. Bellisle F. Meals and snacking, diet quality and energy balance. Physiol Behav. 2014;134:38–43.
59. Stano S, Alam F, Wu L, Dutia R, Ng S-N, Sala M, et al. Effect of meal size and texture on gastric pouch emptying and glucagon-like peptide 1 after gastric bypass surgery. Surg Obes Relat Dis. 2017;13:1975–83.
60. Parkman HP, Camilleri M, Farrugia G, Mccallum RW, Bharucha AE, Mayer EA, et al. Gastroparesis and functional dyspepsia: excerpts from the AGA/ANMS meeting. Neurogastroenterol Motil. 2010;22:113–33.
61. Hinshaw L, Schiavon M, Mallad A, Man CD, Basu R, Bharucha AE, et al. Effects of delayed gastric emptying on postprandial glucose kinetics, insulin sensitivity, and β-cell function. Am J Physiol Endocrinol Metab. 2014;307:E494–502.
62. Speakman JR. The roles of different macronutrients in regulation of appetite, energy intake and adiposity. Curr Opin Endocr Metab Res. 2022;22:100297.
63. Carreiro AL, Dhillon J, Gordon S, Higgins KA, Jacobs AG, McArthur BM, et al. The macronutrients, appetite, and energy intake. Annu Rev Nutr. 2016;36:73–103.
64. Pal S, Radavelli-Bagatini S, Hagger M, Ellis V. Comparative effects of whey and casein proteins on satiety in overweight and obese individuals: a randomized controlled trial. Eur J Clin Nutr. 2014;68:980–6.
65. Pasman WJ, Blokdijk VM, Bertina FM, Hopman WPM, Hendriks HFJ. Effect of two breakfasts, different in carbohydrate composition, on hunger and satiety and mood in healthy men. Int J Obes. 2003;27:663–8.
66. Cooper JA, Stevenson JL, Paton CM. Hunger and satiety responses to saturated fat-rich meals before and after a high PUFA diet. FASEB J. 2016;30:405.7.
67. Leidy HJ, Apolzan JW, Mattes RD, Campbell WW. Food form and portion size affect postprandial appetite sensations and hormonal responses in healthy, nonobese, older adults. Obesity. 2010;18:293–9.
68. Miquel-Kergoat S, Azais-Braesco V, Burton-Freeman B, Hetherington MM. Effects of chewing on appetite, food intake and gut hormones: a systematic review and meta-analysis. Physiol Behav. 2015;151:88–96.
69. Hill AJ, Blundell JE. Nutrients and behaviour: research strategies for the investigation of taste characteristics, food preferences, hunger sensations and eating patterns in man. J Psychiatr Res. 1982;17:203–12.
70. Chaput J-P, Gilbert J-A, Gregersen NT, Pedersen SD, Sjödin AM. Comparison of 150-mm versus 100-mm visual analogue scales in free living adult subjects. Appetite. 2010;54:583–6.

71. Raben A, Agerholm-Larsen L, Flint A, Holst JJ, Astrup A. Meals with similar energy densities but rich in protein, fat, carbohydrate, or alcohol have different effects on energy expenditure and substrate metabolism but not on appetite and energy intake. Am J Clin Nutr. 2003;77:91–100.
72. Gibbons C, Caudwell P, Finlayson G, King N, Blundell J. Validation of a new hand-held electronic data capture method for continuous monitoring of subjective appetite sensations. Int J Behav Nutr Phys Act. 2011;8:57.
73. Zhu Y, Blundell JE, Holschuh NM, McLean R, Menon RS. Validation of a mobile app-based visual analog scale for appetite measurement in the real world: a randomized digital clinical trial. Nutrients. 2023;15:304.
74. Sadoul BC, Schuring EAH, Mela DJ, Peters HPF. The relationship between appetite scores and subsequent energy intake: an analysis based on 23 randomized controlled studies. Appetite. 2014;83:153–9.
75. Reid C, Harbron C, Blundell J, Stubbs J. What are psychometric assessments of appetite asking: a preliminary multivariate analysis. Int J Obes. 1998;22:151.
76. Stunkard AJ, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. J Psychosom Res. 1985;29:71–83.
77. Van Strien T, Frijters J, Bergers G, Defares P. The Dutch eating behavior questionnaire (DEBQ) for assessment of restrained, emotional and external eating behavior. Int J Eat Disord. 1986;5:295–315.
78. Bennett C, Blissett J. Measuring hunger and satiety in primary school children. Validation of a new picture rating scale. Appetite. 2014;78:40–48.
79. Triador L, Colin-Ramirez E, Mackenzie ML, Tomaszewski E, Shah K, Gulayets H, et al. A two-component pictured-based appetite assessment tool is capable of detecting appetite sensations in younger children: a pilot study. Nutr Res. 2021;89:45–55.
80. Jackson RA, Peters N, Advani U, Perry G, Rogers J, Brough WH, et al. Forearm glucose uptake during the oral glucose tolerance test in normal subjects. Diabetes. 1973;22:442–58.
81. Chen Y, Edinburgh RM, Hengist A, Smith HA, Walhin J, Betts JA, et al. Venous blood provides lower glucagon-like peptide-1 concentrations than arterialized blood in the postprandial but not the fasted state: consequences of sampling methods. Exp Physiol. 2018;103:1200–5.
82. Edinburgh RM, Hengist A, Smith HA, Betts JA, Thompson D, Walhin J-P, et al. Prior exercise alters the difference between arterialised and venous glycaemia: implications for blood sampling procedures. Br J Nutr. 2017;117:1414–21.
83. Forster HV, Dempsey JA, Thomson J, Vidruk E, DoPico GA. Estimation of arterial PO2, PCO2, pH, and lactate from arterialized venous blood. J Appl Physiol. 1972;32:134–7.
84. Jensen MD, Heiling VJ. Heated hand vein blood is satisfactory for measurements during free fatty acid kinetic studies. Metabolism. 1991;40:406–9.
85. Liu D, Moberg E, Kollind M, Lins P-E, Adamson U, Macdonald IA. Arterial, arterialized venous, venous and capillary blood glucose measurements in normal man during hyperinsulinaemic euglycaemia and hypoglycaemia. Diabetologia. 1992;35:287–90.
86. Nauck MA, Ließ H, Siegel EG, Niedmann PD, Creutzfeldt W. Critical evaluation of the 'heated-hand-technique' for obtaining 'arterialized' venous blood: incomplete arterialization and alterations in glucagon responses. Clin Physiol. 1992;12:537–52.
87. Watkins J, Simpson A, Betts JA, Thompson D, Holliday A, Deighton K, et al. Galactose ingested with a high-fat beverage increases postprandial lipemia compared with glucose but not fructose ingestion in healthy men. J Nutr. 2020;150:1765–72.
88. Hengist A, Davies RG, Rogers PJ, Brunstrom JM, van Loon LJC, Walhin J-P, et al. Restricting sugar or carbohydrate intake does not impact physical activity level or energy intake over 24 h despite changes in substrate use: a randomised crossover study in healthy men and women. Eur J Nutr. 2022;62:921–40.
89. Chen YC, Walhin JP, Hengist A, Gonzalez JT, Betts JA, Thompson D. Interrupting prolonged sitting with intermittent walking increases postprandial gut hormone responses. Med Sci Sports Exerc. 2022;54:1183–9.
90. Broom DR, Miyashita M, Wasse LK, Pulsford R, King JA, Thackray AE, et al. Acute effect of exercise intensity and duration on acylated ghrelin and hunger in men. J Endocrinol. 2017;232:411–22.
91. Gonzalez JT, Veasey RC, Rumbold PLS, Stevenson EJ. Breakfast and exercise contingently affect postprandial metabolism and energy balance in physically active males. Br J Nutr. 2013;110:721–32.
92. Hagan RD, Diaz FJ, Horvath SM. Plasma volume changes with movement to supine and standing positions. J Appl Physiol. 1978;45:414–7.
93. Holst JJ, Wewer Albrechtsen NJ. Methods and guidelines for measurement of glucagon in plasma. Int J Mol Sci. 2019;20:5416.
94. Rehfeld JF. On measurement of cholecystokinin in plasma with reference to obesity studies. Nutr Res. 2020;76:1–8.
95. Kojima M, Hosoda H, Date Y, Nakazato M, Matsuo H, Kangawa K. Ghrelin is a growth-hormone releasing acylated peptide. Nature. 1999;402:656–60.
96. Hougland JL. Ghrelin octanoylation by ghrelin O-acyltransferase: unique protein biochemistry underlying metabolic signaling. Biochem Soc Trans. 2019;47:169–78.
97. Batterham RL, Cowley MA, Small CJ, Herzog H, Cohen MA, Dakin CL, et al. Gut hormone PYY3-36 physiologically inhibits food intake. Nature. 2002;418:650–4.
98. Delhanty PJD, Huisman M, Julien M, Mouchain K, Brune P, Themmen APN, et al. The acylated (AG) to unacylated (UAG) ghrelin ratio in esterase inhibitor-treated blood is higher than previously described. Clin Endocrinol. 2015;82:142–6.
99. Hosoda H, Kangawa K. Standard sample collections for blood ghrelin measurements. In: Masayasu Kojima, Kenji Kangawa (eds) Methods in enzymology. San Diego, CA, Academic Press Inc.; 2012. pp. 113–26.
100. Deschaine SL, Leggio L. Understanding plasma treatment effect on human acyl-ghrelin concentrations. Eur Rev Med Pharm Sci. 2020;24:1585–9.
101. Blatnik M, Soderstrom CI. A practical guide for the stabilization of acylghrelin in human blood collections. Clin Endocrinol. 2011;74:325–31.
102. Holst JJ. From the incretin concept and the discovery of GLP-1 to today's diabetes therapy. Front Endocrinol. 2019;10:260.
103. Deacon CF, Nauck MA, Meier J, Hücking K, Holst JJ. Degradation of endogenous and exogenous gastric inhibitory polypeptide in healthy and in type 2 diabetic subjects as revealed using a new assay for the intact peptide. J Clin Endocrinol Metab. 2000;85:3575–81.
104. Kuhre RE, Wewer Albrechtsen NJ, Hartmann B, Deacon CF, Holst JJ. Measurement of the incretin hormones: glucagon-like peptide-1 and glucose-dependent insulinotropic peptide. J Diab Complications. 2015;29:445–50.
105. Vilsbøll T, Krarup T, Madsbad S, Holst JJ. Both GLP-1 and GIP are insulinotropic at basal and postprandial glucose levels and contribute nearly equally to the incretin effect of a meal in healthy subjects. Regul Pept. 2003;114:115–21.
106. Deacon CF, Pridal L, Klarskov L, Olesen M, Holst JJ. Glucagon-like peptide 1 undergoes differential tissue-specific metabolism in the anesthetized pig. Am J Physiol Endocrinol Metab. 1996;271:E458–64.
107. Bak MJ, Albrechtsen NW, Pedersen J, Hartmann B, Christensen M, Vilsbøll T, et al. Specificity and sensitivity of commercially available assays for glucagon and oxyntomodulin measurement in humans. Eur J Endocrinol. 2014;170:529–38.
108. Muscogiuri G, DeFronzo RA, Gastaldelli A, Holst JJ. Glucagon-like peptide-1 and the central/peripheral nervous system: crosstalk in diabetes. Trends Endocrinol Metab. 2017;28:88–103.
109. Wewer Albrechtsen NJ, Asmar A, Jensen F, Törang S, Simonsen L, Kuhre RE, et al. A sandwich ELISA for measurement of the primary glucagon-like peptide-1 metabolite. Am J Physiol Endocrinol Metab. 2017;313:E284–91.
110. Ørskov C, Rabenhøj L, Wettergren A, Kofod H, Holst JJ. Tissue and plasma concentrations of amidated and glycine-extended glucagon-like peptide I in humans. Diabetes. 1994;43:535–9.
111. Vilsbøll T, Krarup T, Deacon CF, Madsbad S, Holst JJ. Reduced postprandial concentrations of intact biologically active glucagon-like peptide 1 in type 2 diabetic patients. Diabetes. 2001;50:609–13.
112. Yanagimachi T, Fujita Y, Takeda Y, Honjo J, Sakagami H, Kitsunai H, et al. Dipeptidyl peptidase-4 inhibitor treatment induces a greater increase in plasma levels of bioactive GIP than GLP-1 in non-diabetic subjects. Mol Metab. 2017;6:226–31.
113. Yabe D, Watanabe K, Sugawara K, Kuwata H, Kitamoto Y, Sugizaki K, et al. Comparison of incretin immunoassays with or without plasma extraction: Incretin secretion in Japanese patients with type 2 diabetes. J Diab Investig. 2012;3:70–79.
114. Rasmussen C, Richter MM, Jensen NJ, Heinz N, Hartmann B, Holst JJ, et al. Preanalytical impact on the accuracy of measurements of glucagon, GLP-1 and GIP in clinical trials. Scand J Clin Lab Invest. 2023;83:591–8.
115. Bak MJ, Wewer Albrechtsen NJ, Pedersen J, Knop FK, Vilsbøll T, Jørgensen NB, et al. Specificity and sensitivity of commercially available assays for glucagon-like peptide-1 (GLP -1): implications for GLP -1 measurements in clinical studies. Diab Obes Metab. 2014;16:1155–64.
116. Eberlein GA, Eysselein VE, Schaeffer H, Layer P, Grandt D, Goebell H. A new molecular form of PYY: structural characterization of human PYY(3-36) and PYY(1-36). Peptides. 1989;10:797–803.
117. Dos M, Medeiros S, Turner AJ. Processing and metabolism of peptide-YY: pivotal roles of dipeptidylpeptidase-IV, aminopeptidase-P, and endopeptidase-24.11. Endocrinology. 1994;134:2088–94.
118. Toräng S, Veedfald S, Rosenkilde MM, Hartmann B, Holst JJ. The anorexic hormone peptide yy3-36 is rapidly metabolized to inactive peptide YY3-34 in vivo. Physiol Rep. 2015;3:e12455.
119. Kowalka AM, Alexiadou K, Cuenco J, Clarke RE, Minnion J, Williams EL, et al. The postprandial secretion of peptide YY_{1-36} and $_{3-36}$ in obesity is differentially increased after gastric bypass versus sleeve gastrectomy. Clin Endocrinol. 2023;99:272–84.
120. Karra E, O'Daly OG, Choudhury AI, Yousseif A, Millership S, Neary MT, et al. A link between FTO, ghrelin, and impaired brain food-cue responsivity. J Clin Investig. 2013;123:3539–51.
121. Weissgerber TL, Winham SJ, Heinzen EP, Milin-Lazovic JS, Garcia-Valencia O, Bukumiric Z, et al. Reveal, don't conceal. Circulation. 2019;140:1506–18.

122. Weissgerber TL, Milic NM, Winham SJ, Garovic VD. Beyond bar and line graphs: time for a new data presentation paradigm. PLoS Biol. 2015;13:e1002128.
123. Vickers AJ, Altman DG. Statistics notes: analysing controlled trials with baseline and follow up measurements. BMJ. 2001;323:1123–4.
124. Bland JM, Altman DG. Best (but oft forgotten) practices: testing for treatment effects in randomized trials by separate analyses of changes from baseline in each group is a misleading approach. Am J Clin Nutr. 2015;102:991–4.
125. Matthews JN, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. BMJ. 1990;300:230–5.
126. Wolever T, Jenkins D. The use of the glycemie Index in predicting the blood glucose response to mixed meals. Am J Clin Nutr. 1986;43:167–72.
127. Wolever TMS. Effect of blood sampling schedule and method of calculating the area under the curve on validity and precision of glycaemic index values. Br J Nutr. 2004;91:295–300.
128. Narang BJ, Atkinson G, Gonzalez JT, Betts JA. A tool to explore discrete-time data: the time series response analyser. Int J Sport Nutr Exerc Metab. 2020;30:374–81.
129. Cook JA, Julious SA, Sones W, Hampson LV, Hewitt C, Berlin JA, et al. DELTA2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. Trials. 2018;19:606.
130. Cohen J. Statistical power analysis for the behavioural sciences. 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum Associates; 1989.
131. Hopkins WG, Marshall SW, Batterham AM, Hanin J. Progressive statistics for studies in sports medicine and exercise science. Med Sci Sports Exerc. 2009;41:3–12.
132. Li P, Stuart EA. Best (but oft-forgotten) practices: missing data methods in randomized controlled nutrition trials. Am J Clin Nutr. 2019;109:504–8.
133. Gadbury GL, Coffey CS, Allison DB. Modern statistical methods for handling missing repeated measurements in obesity trial data: beyond LOCF. Obes Rev. 2003;4:175–84.
134. Elobeid MA, Padilla MA, McVie T, Thomas O, Brock DW, Musser B, et al. Missing data in randomized clinical trials for weight loss: scope of the problem, state of the field, and performance of statistical methods. PLoS ONE. 2009;4:e6624.
135. LaPorte GJ, Chauff S, Cammack J, Burton-Freeman B, Krakoff J, Stinson EJ, et al. An algorithm to simulate missing data for mixed meal tolerance test response curves. Am J Clin Nutr. 2024;120:145–52.
136. Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J Exp Soc Psychol. 2013;49:764–6.
137. de Souza RJ, Eisen RB, Perera S, Bantoto B, Bawor M, Dennis BB, et al. Best (but oft-forgotten) practices: sensitivity analyses in randomized controlled trials. Am J Clin Nutr. 2016;103:5–17.
138. Goltz FR, Thackray AE, Atkinson G, Lolli L, King JA, Dorling JL, et al. True interindividual variability exists in postprandial appetite responses in healthy men but is not moderated by the FTO genotype. J Nutr. 2019;149:1159–69.
139. Atkinson G, Batterham AM. True and false interindividual differences in the physiological response to an intervention. Exp Physiol. 2015;100:577–88.
140. Senn S, Rolfe K, Julious SA. Investigating variability in patient response to treatment—a case study from a replicate cross-over study. Stat Methods Med Res. 2011;20:657–66.
141. Gibbons C, Caudwell P, Finlayson G, Webb D-L, Hellström PM, Näslund E, et al. Comparison of postprandial profiles of ghrelin, active GLP-1, and total PYY to meals varying in fat and carbohydrate and their association with hunger and the phases of satiety. J Clin Endocrinol Metab. 2013;98:E847–55.
142. Yeomans M, Bertenshaw E. Appetite and food intake: a human experimental perspective. In: Harris R, Mattes R, editors. Appetite and food intake: behavioural and physiological considerations. CRC Press, Boca Raton, Florida; 2008. pp. 61–80.
143. DiMeglio D, Mattes R. Liquid versus solid carbohydrate: effects on food intake and body weight. Int J Obes. 2000;24:794–800.
144. Yeomans M. Rating changes over the course of meals: what do they tell us about motivation to eat?. Neurosci Biobehav Rev. 2000;24:249–59.
145. Gibbons C, Finlayson G, Caudwell P, Webb D-L, Hellström PM, Näslund E, et al. Postprandial profiles of CCK after high fat and high carbohydrate meals and the relationship to satiety in humans. Peptides. 2016;77:3–8.
146. Brouns F, Bjorck I, Frayn KN, Gibbs AL, Lang V, Slama G, et al. Glycaemic index methodology. Nutr Res Rev. 2005;18:145–71.
147. Anderson GH, Catherine NLA, Woodend DM, Wolever TM. Inverse association between the effect of carbohydrates on blood glucose and subsequent short-term food intake in young men. Am J Clin Nutr. 2002;76:1023–30.
148. Gibbons C, Beaulieu K, Almiron-Roig E, Navas-Carretero S, Martínez JA, O'Hara B, et al. Acute and two-week effects of neotame, stevia rebaudioside M and sucrose-sweetened biscuits on postprandial appetite and endocrine response in adults with overweight/obesity—a randomised crossover trial from the SWEET consortium. EBioMedicine. 2024;102:105005.
149. Gibbons C, Blundell J, Tetens Hoff S, Dahl K, Bauer R, Bækdal T. Effects of oral semaglutide on energy intake, food preference, appetite, control of eating and body weight in subjects with type 2 diabetes. Diab Obes Metab. 2021;23:581–8.
150. Green S, Delargy H, Joanes D, Blundell J. A satiety quotient: a formulation to assess the satiating effect of food. Appetite. 1997;29:291–304.
151. Drapeau V, King N, Hetherington M, Doucet E, Blundell J, Tremblay A. Appetite sensations and satiety quotient: predictors of energy intake and weight loss. Appetite. 2007;48:159–66.
152. Arguin H, Tremblay A, Blundell JE, Després J-P, Richard D, Lamarche B, et al. Impact of a non-restrictive satiating diet on anthropometrics, satiety responsiveness and eating behaviour traits in obese men displaying a high or a low satiety phenotype. Br J Nutr. 2017;118:750–60.
153. Drapeau V, Jacob R, Panahi S, Tremblay A. Effect of energy restriction on eating behavior traits and psychobehavioral factors in the low satiety phenotype. Nutrients. 2019;11:245.
154. Taylor R, Magnusson I, Rothman DL, Cline GW, Caumo A, Cobelli C, et al. Direct assessment of liver glycogen storage by 13C nuclear magnetic resonance spectroscopy and regulation of glucose homeostasis after a mixed meal in normal subjects. J Clin Investig. 1996;97:126–32.
155. Åberg S, Webb D-L, Nordin E, Hellström PM, Landberg R. Postprandial effects of four test meals containing wholegrain rye or refined wheat foods on circulating incretins, ghrelin, glucose, and inflammatory markers. J Nutr. 2025;155:185–96.
156. Samkani A, Skytte MJ, Kandel D, Kjaer S, Astrup A, Deacon CF, et al. A carbohydrate-reduced high-protein diet acutely decreases postprandial and diurnal glucose excursions in type 2 diabetes patients. Br J Nutr. 2018;119:910–7.

ACKNOWLEDGEMENTS

The research was supported by the National Institute for Health and Care Research (NIHR) Leicester Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. FA was supported by the Wellcome Trust [223512/Z/21/Z] through the Leicestershire Healthcare Inequalities Improvement Doctoral Training Programme Fellowship.

AUTHOR CONTRIBUTIONS

The concept for the paper was initiated by JK, FA, GW, DP, and DT. JK, SW and DT wrote the 'Introduction'. LM, MM and DT wrote the 'Familiarisation and standardisation' section. CM, FA and GW wrote the 'Meal stimulus' section. CG, GF, JB, KB, MH and VD wrote the 'Appetite scales' section. JG, DB and LJ write the 'Blood sampling' section. JK, SW and DP wrote the 'Appetite-related peptides' section. AT and DS wrote the 'Data handling' section. JK, SW and AT edited the first draft of the paper. All authors reviewed and approved the final version.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41366-025-01866-7.

Correspondence and requests for materials should be addressed to James A. King.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2025