Special Article

Good practice in food-related neuroimaging

Paul AM Smeets,[1,2] Alain Dagher,[3] Todd A Hare,[4] Stephanie Kullmann,[5] Laura N van der Laan,[6] Russell A Poldrack,[7] Hubert Preissl,[5] Dana Small,[8] Eric Stice,[9] and Maria G Veldhuizen[8]

[1]UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht, NL; [2]Division of Human Nutrition and Health, Wageningen University & Research, Wageningen, The Netherlands; [3]Montreal Neurological Institute, McGill University, Montreal, Canada; [4]Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland; [5]Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Center Munich at the University of Tübingen, German Center for Diabetes Research, Tübingen, Germany; [6]Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, The Netherlands; [7]Department of Psychology, Stanford University, Stanford, CA; [8]Department of Psychiatry, Yale School of Medicine, New Haven, CT; and [9]Oregon Research Institute, Eugene, OR

ABSTRACT

The use of neuroimaging tools, especially functional magnetic resonance imaging, in nutritional research has increased substantially over the past 2 decades. Neuroimaging is a research tool with great potential impact on the field of nutrition, but to achieve that potential, appropriate use of techniques and interpretation of neuroimaging results is necessary. In this article, we present guidelines for good methodological practice in functional magnetic resonance imaging studies and flag specific limitations in the hope of helping researchers to make the most of neuroimaging tools and avoid potential pitfalls. We highlight specific considerations for food-related studies, such as how to adjust statistically for common confounders, like, for example, hunger state, menstrual phase, and BMI, as well as how to optimally match different types of food stimuli. Finally, we summarize current research needs and future directions, such as the use of prospective designs and more realistic paradigms for studying eating behavior. *Am J Clin Nutr* 2019;109:491–503.

Keywords: functional magnetic resonance imaging, neuroimaging, good practice, data sharing, food viewing, food choice, taste, aroma, satiation

Current State of the Field of Nutritional Neuroimaging

Introduction

The brain plays a central role in the regulation of food intake. It integrates many different state- and trait-related neural and hormonal signals that affect eating behavior. Understanding how normal and maladaptive eating behaviors emerge and are maintained is crucial for developing effective eating interventions or treatments, such as weight loss or maintenance programs. Thus, studying the brain structures and processes underlying eating behavior has great potential significance, especially when combined with information on other aspects of physiology and psychology.

Since the late 1990s, functional neuroimaging techniques have been increasingly used to study food-related brain activity in humans. Among the first studies were taste/flavor positron-emission tomography (PET) studies (1) and fMRI (2) and PET studies on the effects of extreme hunger in healthy (3) and obese (4) individuals. Since then, fMRI in particular has become a widely used neuroimaging technique that is often used to study food-related neural correlates in health and disease. We focus here on task-based fMRI, but many of the issues addressed apply similarly to resting-state fMRI, PET, and perfusion fMRI, as well as structural MRI studies.

We present a set of guidelines for good practice in the use of neuroimaging with the hope of helping researchers make the most of these powerful but readily misinterpreted or even misused techniques. We view the establishment of a widely accepted set of guidelines as critical at this point in the development of the field, in part because, although simple visual and motor tasks yield large, robust, and readily replicable brain responses in primary visual and motor cortex, higher-order tasks often produce

This work is supported by the European Union Seventh Framework Programme under Grant Agreement 607310 (Nudge-it) (PAMS, TAH, SK, LNV, HP).

Present address for LNV: Tilburg Center for Cognition and Communication, Tilburg University, Tilburg, The Netherlands.

Supplemental Checklist S1, Supplemental Material S2, and Supplemental Table S1 are available from the "Supplementary data" link in the online posting of the article and from the same link in the online table of contents at https://academic.oup.com/ajcn/.

Address correspondence to PAMS (e-mail: p.smeets@umcutrecht.nl).

Abbreviations used: BOLD, blood oxygen level-dependent; COBIDAS, Committee on Best Practice in Data Analysis and Sharing; DXA, dual-energy X-ray absorptiometry; PET, positron-emission tomography; ROI, region of interest; vmPFC, ventromedial prefrontal cortex; VR, virtual reality.

Received July 25, 2018. Accepted for publication November 5, 2018.

First published online March 5, 2019; doi: https://doi.org/10.1093/ajcn/nqy344.

smaller, more variable responses that are harder to replicate. For example, the most commonly used type of fMRI task in the food domain is the presentation of food images. Meta-analyses have shown that even the brain regions most consistently shown to differentially respond to food compared with nonfood images are significantly active in less than 40% of studies (5). Although brain responses to visual food cues in fasted overweight/obese participants have been found to have relatively good mean-level reproducibility, they had poor within-subject test–retest reliability (6). Another example is the fMRI studies that examined the functional significance of the fat mass and obesity-associated gene FTO. Individuals with the "high-risk" AA FTO variant have been found to show less responsivity to high-calorie food images in a fasted state than "low-risk" TT individuals in reward-related brain regions (7). Also, adults with the AA genotype showed less food cue activation in the prefrontal cortex 30 min after ingesting 75 g of glucose than adults without the AA genotype, but no differences in a fasted state (8). In contrast, individuals with the AA or AT genotypes showed greater responsivity to food (9) and high-calorie food images (10, 11) in reward-related brain areas than "low-risk" TT individuals.

This variability in findings is also due, in part, to divergent characteristics of the individual study designs, highlighting the current scarcity and strong need for direct replication studies. Studies of food stimulus responses and eating behavior differ in many important ways including the structure, timing, and stimuli of the fMRI task; software, strategy, and parameter settings used for processing and statistical analysis of the data; and individual characteristics such as age, gender, and eating-related traits, and state variables such as current hunger level and weight status. In addition, the effect size of food-related brain activation is often modest, and isolating specific effects of interest can be challenging because there are many confounders and interacting factors. For example, in a food-viewing task, caloric content may well covary with palatability, and so responses to high- compared with low-calorie foods cannot be attributed to caloric content per se. Further, there are clear individual differences in food preferences and familiarity that introduce additional variance (12). Thus, there is a need for better standardization of the food stimuli and fMRI task designs used and the additional data that are collected on participant's state (hunger, mood) and personal characteristics that may be used to control for confounding effects in the analyses.

In addition to the variability between studies and infrequent replication attempts, a lack of sufficient power and rigor in individual experiments is a key factor. Just as in other fields investigating higher cognitive processes, many of the earlier fMRI studies on eating behavior are underpowered (13, 14). Although there is a clear trend toward larger sample sizes in fMRI over the past decade, only recently have tools for better power calculation become available (15, 16). The need for informed study planning is further highlighted by recent empirical demonstrations stressing the importance of appropriate, validated statistical thresholding approaches (17).

Despite previous shortcomings, there is reason to be optimistic that this situation will improve in the near term. This optimism stems from the ongoing development of neuroimaging hardware and analysis software, and especially the adoption of higher-quality standards in the field. We believe that replication studies and open data sharing will play a central role in the ongoing efforts to advance the utility and reliability of food-related neuroimaging findings. The current lack of replication efforts means that it remains unknown how robust many of the original findings in the field are, and although meta-analyses can give some initial indications, the accuracy of meta-analytic studies is limited by the number and quality of the primary studies they aggregate over and is reduced by publication bias and lack of access to primary data (14). The aim of this article is to foster good practice in food-related neuroimaging by presenting guidelines for good methodological practice, outlining potential pitfalls, and providing recommendations for food-related fMRI task implementation.

What can we learn from fMRI?

fMRI usually refers to blood oxygen level-dependent (BOLD) fMRI. This popular form of fMRI exploits the fact that at a site of increased neuronal firing (brain activation), increased local blood flow leads to a decreased concentration of deoxygenated hemoglobin in the capillaries. This reduces the local distortion of the magnetic field by the para-magnetic deoxy-hemoglobin, which leads to a small increase in the fMRI signal (~0.5–4%). Thus, BOLD fMRI provides an indirect vascular measure of (changes in) neuronal activity. Most fMRI studies use cognitive or sensory tasks in which different task conditions are contrasted to assess neural activation differences of interest (e.g., viewing food images compared with viewing nonfood images or tasting a chocolate milkshake compared with a control solution). This provides information on which brain regions become more or less active during a certain task (functional localization) and whether this differs between study conditions such as hunger and satiety or different groups of participants.

In recent years, there has been increasing focus on (differences in) functional connectivity; that is, the degree to which task-related brain activation in a specific brain region covaries with activation in other brain regions (functional interactions) (18). Also, "resting-state" fMRI, which examines the spatio-temporal networks of correlated activity in the absence of a specific task (lying still with eyes closed, or mere visual fixation), has become a popular and promising means of assessing individual differences in neurobiology (19, 20).

Brain findings per se can be useful, but often their combination with other measures creates synergy and aids the interpretation of fMRI findings; fMRI results become more meaningful when associations with physiological signals and subjective ratings or individual characteristics can be established and when they are linked to relevant outcomes such as food intake (21, 22) and weight change (23–27). Because the brain is so central in the regulation of food intake and body weight, fMRI is well suited for connecting different levels of understanding.

Many brain-imaging studies of neural response to food stimuli seek to make inferences regarding the role of neural responsivity in the development of adverse physical or mental health problems such as obesity or eating disorders. For instance, it had originally been suggested, based on the evidence that obese compared with lean individuals have lower D2 receptor binding as measured by PET, that low responsivity of reward circuitry increases the risk of overeating and consequential obesity (28, 29). However, this is an example of the complexity involved in drawing inferences from cross-sectional studies because they are unable to differentiate

neural vulnerability factors from neural consequences of these physical and mental health problems.

Prospective studies that can show that the putative neural vulnerability factor pre-dates and predicts future emergence of the adverse public health outcome permit stronger inferences than cross-sectional studies. However, they do not rule out the possibility that some omitted third variable explains both the neural response and the emergence of the public health outcome. Indeed, a larger study spanning the full adult age range concluded that there was no relation between D2 receptor levels and body mass index (BMI) in young adults and a positive relation in older individuals (30), casting doubt on the reward deficiency interpretation. Furthermore, a recent meta-analysis failed to find support for the reward-deficiency interpretation as well (31). Together, this work highlights the importance of prospective studies, meta-analysis, and replication in establishing reliable links between brain structure or function and eating behavior or health outcomes.

Prospective neuroimaging studies in the domain of eating behavior can vary in their breadth and duration. The most basic prospective design is to assess neural responses to experimentally manipulated stimuli or measures of brain morphometry at baseline and then test whether individual differences in these variables predict future increases in, or onset of, the health issue of interest (e.g., future weight gain or onset of obesity among initially nonobese participants). Prospective designs that include repeated measurements of neural responses at multiple time points provide information on biological and behavioral trajectories that can capture behavioral and neural plasticity that occurs in response to weight gain or weight loss over time (or vice versa with behavioral or neural interventions). Prospective repeated-measures neuroimaging studies of food-related behavior and health are thus useful for studying the mechanisms of action for prevention and treatment interventions.

Overall, neuroimaging has exciting potential to contribute to our understanding of the causes of obesity. The significant increase in the incidence of obesity over the past 50 y has been attributed to an interaction of individual vulnerability and an obesogenic environment replete with inexpensive high-calorie foods (32). Considerable evidence suggests that substantial individual vulnerability to this obesogenic environment resides in the brain. As in the mental health literature (33), the search for endophenotypes—that is, neural, cognitive, or personality measures that correlate with weight gain and BMI—has the potential to *1*) provide intermediate measures for gene discovery and *2*) provide explanatory mechanisms for the neural computations that lead to overeating and thus potentially inform the development of therapies. Moreover, the combination of endophenotype research and genetics, performed in different age groups, may allow us to disentangle the two-way relation between body mass composition and the brain, because it is known that visceral obesity itself also causes brain changes (34), which may favor further weight gain. However, as with any measurement technique, the ultimate utility of MRI and other neuroimaging methods depends directly on the experimental designs and analysis strategies it is combined with. In the subsequent sections, we highlight the importance of, and aim to provide initial guidance on, good practice and minimal standards in neuroimaging research, with a particular focus on its application to questions surrounding dietary behavior, nutrition, and obesity.

Methodological Aspects—Good Practice and Minimal Standards

Good practice guidelines

A carefully compiled and commonly agreed upon set of good practice guidelines is essential for maximizing the utility of the complex and ever-growing set of neuroimaging techniques available to researchers. Such guidelines facilitate the design, execution, and interpretation of original research studies, and, moreover, allow for testing of reproducibility, accurate replication (13, 35) and better meta-analyses. In light of the need for such guidelines, the Organization for Human Brain Mapping initiated the Committee on Best Practice in Data Analysis and Sharing (COBIDAS), which set out to define best practices for data analysis and results reporting as well as algorithm and data sharing to promote transparency, reliability, and collaboration. This resulted in a position paper (36) and the COBIDAS report (http://biorxiv.org/content/early/2016/05/20/054262), which provides details for proper reporting and specific good practices.

Two of the most important issues for any fMRI study are *1*) power in terms of both the number of participants included and the task design (e.g., number of trials per condition) and *2*) the threshold used for assessing statistical significance and how that was determined, appropriately controlling for multiple comparisons. These comparisons include the testing of multiple voxels and/or regions of interest, but also extend to tests of neuroimaging measures against multiple measures of individual differences in cognition or health status. The following sections will cover multiple aspects of how these general guidelines can be applied to neuroimaging studies of dietary behavior, nutrition, and obesity. An overview of requirements and recommendations is provided in **Table 1**. After briefly summarizing general good practice guidelines for neuroimaging, we discuss specific experimental design and analysis features for studies using visual, olfactory, or physical foods/liquids as stimuli. We would like to note that the AJCN is committed to the COBIDAS standard and encourages authors to follow the recommendations of that report. Upon submission, authors will be asked to complete a checklist based on Appendix D of the COBIDAS report. All items flagged as mandatory need to be satisfied as a minimal standard. This checklist is available as **Supplemental Checklist S1**.

Power calculation and study planning

The prevalence of underpowered studies in neuroimaging, as well as many other scientific disciplines, is one of the biggest, but also most concretely addressable, issues we face (14, 37, 38). Power analysis is important not simply to avoid performing a futile study, but also to ensure that any positive findings are likely to be true positives; as noted in a previous study (37), low power increases the likelihood that any positive findings are false positives and thus reduces the likelihood that findings from underpowered studies are replicable. To date, sample-size calculations based on realistic power analyses have been made only rarely during the planning stages of

TABLE 1 Overview of requirements and recommendations for nutritional neuroimaging[1]

Requirement/recommendation	Level
Participant description	
Report age	M
Report gender and test for possible effects	M
Report race and ethnicity	R
Report handedness and account for nonrighthandedness in analyses	M
Report socio-economic status	R
Report physical activity level	R
Report use of relevant medication, tobacco, alcohol, and caffeine	R
Report menstrual cycle phase and how this was accounted for in the analysis	HR
Report BMI or age-adjusted BMI and test for possible effects	M
Report further adiposity measures, e.g., percentage body fat, waist–hip ratio	R
Report a measure of dietary restraint	R
Report a measure of stress	R
Report personality traits such as reward sensitivity and impulsivity	R
Eating disorder scales	R
Report weight history; weight lost or gained in the weeks before brain imaging	HR
Report time since last meal	M
Standardize the last meal before brain imaging	R
Report appetite ratings	HR
Report thirst ratings	R
Study design/procedures	
Describe the hunger state(s) and how they were achieved	M
Report food stimulus details including macronutient composition and energy content	M
For pre- compared with postfeeding studies, motivate why fasted and fed conditions could not be completed on separate days to avoid order effects	M
fMRI task	
Mandatory items in the COBIDAS checklist (S1)	M
Provide a power calculation	HR
Report the task instructions	M
Report the number and timing of the task events and how their order was randomized and/or optimized	M
Describe the stimuli used and how they were matched, e.g., on visual characteristics	M
Report stimulus liking and where appropriate intensity	M
For taste stimuli: report temperature, volume, flow rate, and swallowing instructions	M
For olfactory stimuli: report temperature, flow rate, and sniffing instructions	M
fMRI data analysis	
Mandatory items in the COBIDAS checklist (S1)	M
Indicate how correction for multiple comparisons was done and how the threshold used was determined	M
Test multiple ROIs with a single combined ROI mask	M
Use appropriate covariates, such as stimulus liking, gender, menstrual cycle phase, and BMI	HR
Include blood parameters as covariates, if available	R
Statistical inference/interpretation	
Avoid reverse inference	HR
Be as specific as possible in the degree of overlap when comparing activated brain regions with regions found in other studies	HR

[1] General requirements and recommendations for reporting neuroimaging methods can be found in the COBIDAS checklist (Supplemental Checklist S1).
COBIDAS, Committee on Best Practice in Data Analysis and Sharing; HR, highly recommended; M, mandatory; R, recommended; ROI, region of interest.

fMRI studies. At least, such calculations are rarely reported in literature. This shortcoming is by no means specific to the use of fMRI for nutrition research, but is nonetheless a serious limitation and often results in inconclusive, nonreplicable, or even misleading findings. We now know that common rules of thumb about statistical power for fMRI studies (e.g., 20–30 participants per group) do not hold in many cases and often result in underpowered studies, particularly when the goal is to examine individual differences (39). Underpowered studies are most often a waste of funding as well as the time and effort of both researchers and study participants (38, 40).

Making realistic power calculations requires careful thought and effort, but the necessary tools for doing so are available. Most statistical software packages include dedicated functions for power analyses. Moreover, in recent years, more accessible and fMRI-specific tools [e.g., (15, 41)] have been developed to help researchers make appropriate power calculations that incorporate both within- and between-subjects factors. It is important to remember that power is a function of the number of participants but also of the heterogeneity of the study population and the amount and quality of data collected per participant. In conjunction with sample-size calculations, it is important to optimize the design of fMRI tasks in terms of the number,

temporal distribution, and duration of different trial types (for general guidelines, see http://imaging.mrc-cbu.cam.ac.uk/imaging/DesignEfficiency; for an example of a tool for testing efficiency of an fMRI task design, see http://www.neuropowertools.org/).

The ever-growing number of studies in the literature and the move toward open data sharing mean that, in many cases, data are readily available for use in making estimates of power and requisite sample sizes for new studies. However, it should be noted that effect sizes based only on published studies are likely to be inflated due to publication bias. Therefore, the use of existing data should generally be complemented by piloting the exact experimental procedures. In many cases, researchers and funding agencies will still need to invest significant time and resources into collecting more specific pilot data to make realistic power calculations. However, the returns on such initial investments are worthwhile, and the cost of not conducting appropriate study planning is far greater.

Last, we note that collecting more data (trials or subjects) is not the only way to improve statistical power in fMRI research. The traditional method for analyzing fMRI data (i.e., the mass univariate approach) involves the repeated testing of a regression model in tens or hundreds of thousands of individual voxels. These multiple tests require corrections for multiple comparisons that reduce statistical power. These corrections are necessary for valid inference and cannot be avoided for mass univariate analyses. However, mass univariate analyses are only one means of analyzing fMRI data (42). Multivariate analyses (43) and data-reduction or aggregation techniques such as independent or principal-components analyses, or predefined regions of interest (ROIs), significantly reduce the number of comparisons conducted and thus the degree of correction for multiple comparisons required. Beyond simply increasing power, there is ample reason to believe that multivariate and network-level analyses (44, 45) provide additional insight into brain function, and the application of such techniques to the domain of food choice and nutrition represents an important and as yet relatively underexploited opportunity.

Proper experimental and task design

Eating behavior and nutritional decisions are determined by a plethora of factors. In order to draw strong conclusions from neuroimaging results, we have to know precisely which factors were controlled and which were manipulated. The nature of the scientific question will determine exactly which geno- and phenotypic information is most appropriate to measure or manipulate and report. It is now standard to report BMI as an anthropometric measure, age of the participants, and sex. However, for many specific questions, a deeper phenotyping may be necessary. For example, it is clear that BMI does not provide enough information concerning body composition (46). Better methods to describe the body composition are bioimpedance measures, dual-energy X-ray absorptiometry (DXA), MRI, or BOD POD assessment of body composition. However, the method used for a given study should be appropriate for the aims of the study and justified in terms of costs and benefits to both researchers and participants. Ideally, however, there should be overlap in the measures used to allow better accumulation of evidence. Accordingly, a set of high-priority measures, including MRI, has been proposed to achieve common usage and thereby increase the breadth and impact of obesity research (47).

Hunger state and related factors.

An important factor to control in nutritional studies is hunger state and caloric deprivation because they affect food wanting and food-related brain responses (3, 48–51). In addition, the quantification of food intake is especially important for intervention studies because nutritional composition can also affect neuronal processes. For example, fasting-state studies generally require a 12-h fast and try to control for the subjective hunger state using visual analog scale measures of appetite. However, it has been established that the macronutrient composition of even a single meal can affect hormonal responses extending beyond 12 h (52). Thus, there is added value in the assessment, and inclusion as covariates in analyses, of major hormonal factors related to nutrition. For example, glucose, insulin, leptin, and ghrelin could be included for nutritional studies of neural responses in specifically induced feeding states such as hunger compared with satiety. This would allow researchers to disentangle physiological and subjective factors related to eating processes.

Another issue is that nutritional preferences are culturally and individually determined, and therefore the creation and use of standardized food stimuli can be difficult. Moreover, these evaluations are dependent on the time of day, season, and (hunger) state. For example, a heavy breakfast with savory components is very uncommon in many parts of the world, and if studies are performed during the morning hours this has to be taken into account. Thus, acquiring individual evaluations of the experimental stimuli is another standard operating procedure that should be incorporated into neuroimaging studies of nutrition-related behavioral or physiological responses. In addition, it is advisable to use a standardized meal, for example on the evening before the measurement, or at least to request participants in a repeated-measurements design to consume the same meal preceding all measurements.

Finally, an important challenge in all nutritional studies, including those using neuroimaging, is that the assessment of nutritional intake is difficult to quantify in normal daily life. Currently, most studies use diaries for nutritional intake. However, such self-reports are unreliable (53). There are several ongoing efforts to measure nutritional intake using smartphone applications. However, an assessment of the validity and degree of advantage or disadvantage of smartphone-based methods relative to traditional diary methods and the doubly labeled water method for assessing habitual caloric intake will require further study.

Personal characteristics.

In addition to physiological factors, care must be taken to account and, whenever possible, control for psychological factors in studies of the neurobiology of eating behavior. Personality or cognitive traits may modulate food-related brain responses (12).

Most studies test for eating disorders to exclude clinically relevant diseases. However, it would be advisable to control statistically for subclinical scores on eating-disorder scales.

Choosing and matching food-related stimuli.

Eating engages all of our senses. The extraoral sensations of vision and olfaction provide information about food availability to guide food acquisition. The oral sensations of somatosensation (e.g., texture and temperature), chemesthesis (e.g., astringency, spiciness), gustation (sweet, sour, salty, bitter, umami, and possibly fat and starch taste), and retronasal olfaction provide information to guide consumption once the food is acquired and in the mouth. For example, one uses oral somatosensation to localize a bone in a bite of fish that needs to be extracted before swallowing, whereas the taste of sweetness produces a metabolic cascade to facilitate glucose metabolism (54). The choice of stimulus will depend upon the particular goals of the study. An in-depth discussion of relevant factors to consider for visual, olfactory, and oral food-related stimulation is provided as **Supplemental Material S2**.

fMRI data analysis

Statistical thresholding for whole-brain and regions of interest.

fMRI analyses can be performed at several levels. When using common mass univariate approaches that take all voxels in the brain into account, appropriate corrections for multiple comparisons must be implemented. This has been noted early on (55) but was highlighted several years ago by a conference paper reporting on scans of a dead salmon that was instructed to perform an emotion-recognition task (56). When appropriate correction techniques were not applied, there appeared to be task-related brain activation in the salmon. Naturally, these false-positive activations were no longer seen when appropriate corrections for multiple testing were used.

This infamous "case study" is a salient reminder of the importance of using appropriate statistical methodology in the analysis of neuroimaging data. In many subfields of neuroimaging, it has been commonplace to use rule-of-thumb corrections for multiple comparisons (e.g., a voxel-level threshold of $P < 0.001$ uncorrected combined with a cluster-extent threshold of 10 voxels). However, it is now clear from creative examples such as the salmon study and more rigorous and extensive investigations that such rules are inadequate in controlling false-positive rates. Recent comparisons of correction methods for multiple testing in fMRI data indicate that permutation-based procedures are the best choice and that cluster-based methods should be used correctly (17, 57). Specifically, when Gaussian random field theory is used for cluster-based inference, the cluster-forming threshold should be $P = 0.001$ to avoid inflated false-positive rates (17). More stringent cluster-forming thresholds also help to avoid problems in interpreting the very large activation clusters that often result from low cluster-forming thresholds (57). Note that cluster-based corrected findings indicate that there is likely to be significant activity somewhere within the cluster rather than indicating that all voxels within the cluster are significant. Thus, if we only show that somewhere within a very large cluster there is probably a significant difference between conditions or groups, then we cannot infer or conclude much at all.

In addition to whole-brain analyses, the current literature on the neurobiology of nutrition is substantial enough to justify ROI analysis for certain brain regions or connections between regions. However, in order to be valid, ROI analyses must be planned a priori, ideally preregistered, and the hypotheses about the region must be clearly stated. To avoid biased results, both anatomic and functional ROIs should be defined based on *independent* data sets or functional localizer tasks. Note that multiple comparison corrections must be applied across the ROIs when multiple ROIs are tested for a given hypothesis. Furthermore, the assumptions underlying cluster-based correction methods are rarely satisfied in small volume analyses, and their use in this case should be avoided (58).

Minimizing the influence of movement.

fMRI data are prone to movement-related artifacts because movement causes displacement and distortions in the data. In particular, oral stimulation can be accompanied by significant movement. Movement from swallowing and other activities such as breathing may be larger because of greater body mass. In addition, there is evidence that head motion and BMI share genetic influences, suggesting that movement is a neurobehavioral trait that is greater in obesity [e.g., (59)]. These movements can be counteracted in real time or modeled post hoc during data analysis.

a) Real time. Movement can be minimized physically by the use of cushions around the head, a personalized head case from a 2-part foam mold, or a bite bar. Movement can also be minimized through behavioral training or feedback. One way is to provide the participant with a stationary reference, which has been done by using a cloth strap or tape across the forehead that attaches to the head coil. When the participant moves, they can clearly feel this by the friction on their forehead. This feedback works well and leads to substantial improvement because movement from swallowing mostly results in small movement in the z-plane, which is hard to feel in most head coils. Again, training is important to improve comfort and ability to lie still. Training will also allow participants to learn to swallow with minimal movement of the head by isolating movement to the jaw and tongue during swallowing. The use of real-time feedback with a head motion tracker in a mock scanner may be most efficient (available, for example, at Psychology Software Tools https://pstnet.com/). Scanners with newer software may include real-time monitoring of movement and allow experimenters to immediately redo runs that invoked too much movement. Another solution is to remove the need to swallow altogether by suctioning out liquids (60) or instructing participants to hold the liquid in their mouth until they receive a cue to swallow (61). The downside of these methods, as elaborated in Supplemental Material S2, is that a large area of stimulation is overlooked, that aromas in flavors cannot be perceived, and that an important part of the process of ingestion is omitted.
b) Post hoc analysis. Correction for head motion via image registration is performed as a standard part of the fMRI preprocessing pipeline, but it is clear that this is not sufficient to remove the residual effects of head motion on image intensities (62); for this reason, motion parameters and their derivatives (which quantify change from time point to time point) are often included as nuisance regressors in the

statistical model. However, these too may be insufficient to address large amounts of motion, and it is common to reject data from individual participants, runs, or time points based on motion estimates. The state-of-the-art techniques for motion detection and cleaning have been developed in the context of resting-state fMRI, where head motion is a critical problem (63). In addition to the use of motion estimates and their derivatives as nuisance regressors, it is common to compute a measure of "frame wise displacement," which measures the overall displacement of the images between each pair of subsequent time points, and a measure called DVARS (derivative or root mean square variance over voxels), which quantifies the mean change in image intensity between time points. These measures may be used to "scrub" time points with motion that exceeds a particular threshold, varying from 0.2 to 0.5 mm frame-wise displacement (64), along with surrounding time points; in the context of task-based fMRI analysis, this scrubbing can be performed as part of the statistical model by including single time point regressors for each excluded time point in the model (65). Individual runs or subjects exceeding a threshold level of scrubbed volumes may be dropped; the use of faster imaging with multislice acquisition can improve the handling of motion by reducing the relative amount of data that needs to be removed.

An estimate of vigor of swallowing and exact timing of swallowing may be obtained with expanding bellows and a spirometer (66, 67), which will allow the use of swallowing as either the onset of an event of interest or, alternatively, a nuisance regressor to be covaried out. Similarly, movement from breathing can be estimated with most standard scanner equipment and incorporated into the single-subjects analysis. These variables can also be included as regressors in group analyses to address their confounding effects. Finally, independent component analysis can be used to remove the effects of motion artifacts and physiological noise from breathing and heart beating (68–70).

Analysis of prospective designs.

Although significant advances have occurred in analytic approaches for longitudinal data that better account for autocorrelation of data from the same participant over time, missing data, and nested data (71), these advances are not supported by commonly used fMRI analytic packages. The most basic approach if the data are only collected at 2 time points is to use change scores for the outcome (e.g., T2 BMI–T1 BMI) and simply regress the change scores on BOLD response from the contrast of interest [e.g., (72)]. However, it is critical to covary for baseline BMI because a change in an outcome over time is typically negatively related to baseline values of the outcome (73). Ideally, we recommend using random effects growth mixture models or other types of hierarchical linear models that use full information maximum likelihood to confirm that we model change in behavioral outcomes optimally. This is particularly important when data are collected at 3 or more time points, because there is the potential for nonlinear change over time (e.g., quadratic growth). The slopes and intercepts (coded to reflect baseline values) can then be exported to any of the standard fMRI analytic statistical packages and the slopes regressed against the BOLD response, controlling for the intercept [e.g., (27)]. For repeated-measures studies, which can include natural history observational studies [e.g., (74)] or intervention trials [e.g., (75)], one can simply use repeated-measures ANOVA models to test for differential change in BOLD response in contrasts of interest over time across 2 or more groups. Although one might be tempted to directly contrast BOLD response to the event of interest (e.g., taste of milkshake) from multiple assessment points, we do not recommend this approach because a number of factors can contribute to a variation in BOLD signal over time (e.g., variability in physiological variables, instability of MRI hardware), which may introduce bias. Instead, the contrast of the event of interest against an appropriate control event (e.g., tasting tasteless control solution) should be used. An alternative approach is to read out parameter estimates from the contrast of interest at each assessment and use standard data analytic packages, such as SAS or R to conduct regression models or repeated-measures analyses, but this requires an ROI approach, which does not make use of all the data collected and may miss important peaks that were not anticipated a priori.

Predictive modeling.

One of the potential uses of MRI is the prediction of future outcomes, such as eating behavior, weight change, or treatment responses. A mounting number of studies suggest that neural food cue reactivity can predict outcomes such as energy intake outside the laboratory (76), weight gain (27, 77, 78), weight variability (79), and weight-loss success (23, 80).

However, care must be used during model fitting in order to achieve predictive accuracy on new samples. When model-fit and goodness-of-fit estimates are obtained from the same data, the estimated goodness of fit is inflated because the data have in a sense been used twice (81). One approach to address this is to use cross-validation to assess out-of-sample predictive accuracy; in this method, the model is fit iteratively to subsets of the data and tested on the remaining data that were held out during training (https://web.stanford.edu/~hastie/ElemStatLearn/). This method provides more accurate estimates of how well the model can predict outcomes in new samples; however, predictive accuracies can be highly variable with small samples (82), and accuracies can be inflated if many different parameter sets are tested without proper control (83). For this reason, testing a model (e.g., regression, support vector machine, etc.) fit to one data set against an entirely separate and independent data set remains the gold standard for quantification of predictive accuracy.

Preregistration and data sharing

The importance of transparency for reproducible research is increasingly realized. Studies can be registered at accredited public trial registries such as clinicaltrials.gov, but that does not preclude exploration of the data beyond the testing of the primary hypotheses, although study plans including planned analyses can also be preregistered [e.g., at the Open Science Framework (osf.io)]. To counter publication bias, an increasing number of journals accept registered reports; the study plan is peer-reviewed, and, if accepted, the journal will publish the

results of the planned analyses regardless of their nature (see https://cos.io/rr/).

Transparency and reproducibility are further aided by the sharing of research materials such as task scripts and analysis code as well as the data. There is a spectrum of data sharing, which involves a trade-off between the ease of sharing and the utility of the data (84). On the one hand, meta-analysis has largely relied upon activation coordinates from published papers (85, 86), which are easy to obtain but limited in comparison with meta-analysis based on full statistical images (87). For this reason, it is now recommended to share the unthresholded statistical images from neuroimaging studies using a database such as Neurovault (88). At the other end of the spectrum is the sharing of complete raw data sets via resources such as OpenNeuro, INDI/FCP, and NITRC. The sharing of raw data sets requires substantially more time and effort than sharing of coordinates or statistical results but provides greater utility of the data, such as allowing different analyses to be applied to the same data or allowing raw data to be combined across studies in a "mega-analysis." Recent projects such as the Human Connectome Project (89) and ENIGMA Consortium (90) have demonstrated the substantial utility of sharing large samples of raw MRI data.

Appropriate Interpretation

What can be concluded from fMRI findings (and what cannot)?

Although research on the exact meaning of changes in the BOLD fMRI signal is still ongoing, most researchers assume that differences in BOLD signal reflect differences in neuronal activity "averaged" over the piece of brain tissue that was sampled (voxel). One could argue that as long as we can detect apparently meaningful differences between conditions or groups, BOLD fMRI is of use regardless of the exact underlying neuronal and physiological correlates of these signal differences. Nevertheless, underlying processes such as coupling between neuronal and vascular response may differ between subjects and may be affected by disease states. Notably, obesity is associated with increased cerebrovascular disease risk, and this may affect neurovascular coupling (91). Studies examining cerebrovascular reactivity can be used to assess whether this might be a problem in specific study populations.

A particular point of attention for clinical and intervention studies is that baseline or "resting-state" brain activity may differ between patients and controls or may change due to the study treatment (e.g., meal ingestion or a diet intervention). This may explain observed differences in task-related brain activation, which is usually the main outcome parameter. In addition, because fMRI results usually rely on a comparison between 2 task conditions or groups, the direction of the underlying BOLD signal changes should be examined by extracting cluster parameter estimates to aid interpretation. This allows one to distinguish less deactivation from greater activation, for example. Group × task condition interactions should be reported only where there is a main effect of the task in one of the groups. For example, when there is no clear activation in a region for "food compared with nonfood," great caution should be exercised in reporting and interpreting a group × stimulus type interaction in this area.

It can be challenging to design an fMRI task such that a specific cognitive process is subtracted out by contrasting a task of interest with a control condition. First, in the food domain in particular, it is inherently harder to match stimuli due to their sensory complexity and possible cognitive associations, and we can only approximate control conditions by matching on as many characteristics as we can. Second, the observed differences in regional brain activation may be driven by associated but not necessarily food-specific processes such as arousal, attention, emotion, or motivation. This is not necessarily a drawback, but it is important to be aware of this. Third, fMRI is sensitive such that task instructions and mind set or attentional focus can alter the pattern of brain activation observed [see, e.g., (92–95)]. Thus, when interpreting findings and comparing with the literature, it is important to take seemingly minor differences in task design and instruction into account.

As alluded to before, conclusions can be strengthened by showing that differences in BOLD signal changes correlate with relevant parameters such as stimulus or personal characteristics.

Reverse inference

A common practice in the interpretation of neuroimaging results is the use of reverse inference (96). This refers to interpreting activation of a particular brain region as evidence for the engagement of a particular cognitive process. Although they can provide some information, such inferences are not deductively valid and need further substantiation. In particular, when activation of a brain region cannot be pinpointed to a specific process or when evidence for selective engagement of that region during a specific neural process is weak, reverse inference should be done with caution. For example, areas that are often found to be activated in many studies, also outside the food domain, are the insula, cerebellum, and prefrontal cortex (97, 98). For such large and heterogeneous regions, special care should be taken to consider the exact subregion found in combination with the process of interest. In conclusion, reverse inference should be used with caution and involve as much specificity as possible.

Comparability of findings in "the same" brain region

In general, the discussion of fMRI findings often lacks accuracy. Often it is unclear whether the area being discussed is really in the same part of the larger structure, say within a 10-mm radius, and located in the same hemisphere. This may be particularly true for large areas such as the insula and long gyri (e.g., the inferior frontal gyrus). It is advisable to be as specific as possible (e.g., by distinguishing between anterior, middle, and posterior insula). Likewise, indicative labels such as "dorsolateral prefrontal cortex" or "ventromedial prefrontal cortex" may be used to refer to very different locations. Thus, in all cases, comparison of findings between studies should not be done without checking the exact location to allow appropriate wording of the degree of similarity. In addition, it is important to be clear on the paradigm or other relevant aspects of the study such as the sample size or population used, which can significantly affect comparability of findings and thus the strength of the inferences made. We see the open sharing of unthresholded

group-level statistical maps (e.g., through Neurovault.org) as the most promising way to resolve such ambiguities. If these data are available for all published studies, then comparing the spatial locations of new and existing findings becomes as simple as overlaying 2 or more maps.

A useful approach to overcome regional/functional imprecision is to use meta-analytical results to pinpoint functional areas. Online repositories of meta-analyses such as the ANIMA database (99) or Neurosynth (www.neurosynth.org) (86) can be queried to identify specific functional locations (e.g., the ventromedial prefrontal cortex area that encodes stimulus value or the insular subregion that responds to taste stimuli).

Research Needs and Future Directions

Fostering comparability, data pooling, and meta-analysis

Scientific progress can be promoted by better comparability of research findings, allowing better data pooling and more accurate meta-analyses. This requires better standardization of (neuroimaging) methods and associated measures, along with the application of advanced analysis and modeling techniques to nutritional neuroscience data (100). This would be aided by minimal standards in the field as to which descriptive data must be reported, in addition to common descriptives such as age and gender. This might include as a minimum handedness, BMI, and a measure of hunger state, but could be expanded for many studies by additional measures such as information on diet, body composition (percentage body fat), hormonal status (menstrual cycle phase, appetite-related hormones), and personal as well as personality characteristics (dietary restraint, food attitudes, reward sensitivity, impulsivity).

Task-related fMRI studies would do well to use established paradigms with standardized stimuli adjusted for the population under study and also evaluated by the study participants to confirm (e.g., familiarity). This is aided by sharing of the stimuli used in online databases (see **Supplemental Table S1** in Supplemental Material S2) and sharing of the associated task paradigms and code, preferably at established repositories like the Open Science Framework (https://osf.io/) and GitHub.

An excellent way to make more use of existing data or achieve greater yield from studies is to use the same paradigm and analysis pipeline across many centers. This is particularly useful when it concerns specific (clinical) populations that may be hard to recruit in sufficient numbers by a single center. An example of this is the ENIGMA (90) working groups that assess cortical thickness for different disorders by pooling results obtained from the analysis of anatomic MRI scans from many centers (http://enigma.ini.usc.edu/). Although mainly focused on brain disorders so far, there is an eating disorder group as well (http://enigma.ini.usc.edu/ongoing/enigma-anorexia/).

Another noteworthy initiative is the use of standardized analysis pipelines for neuroimaging data analysis (101) as provided at the OpenNeuro platform (https://www.openneuro.org/). This may help to reduce variation in study results and allows researchers to see how robust their outcomes are when assessed with different software packages. As a minimum, (neuroimaging) analysis scripts should be shared alongside data to better allow replication by others.

Toward predicting future outcomes

The vast majority of nutritional neuroimaging studies are cross-sectional. As alluded to above, to learn more about the causality of obesity and eating disorders it is crucial to promote long-term follow-up studies (e.g., by adding MRI measures to adequately powered cohort studies). Adding to existing or newly formed cohorts would also ensure detailed phenotyping. Individual differences in fMRI task responses or structural data at baseline can then be used to predict future changes in relevant outcomes such as onset of a disease state or growth in symptoms [see, e.g., (27)]. Ideally, phenotyping including neural measures would be done repeatedly to be able to examine neural plasticity that may occur in response to (nutritional) interventions or disease conditions (e.g., onset of an eating disorder or obesity).

Technological advances

More realistic food cue exposure and choice context—potential of virtual reality.

Another direction for future work is the development of more realistic fMRI paradigms which better reflect the reality of food cue exposure and choice. A supine-positioned, immobile participant lying in a narrow, noisy MRI tube, located in a hospital, might reasonably be expected to behave differently than one walking around a supermarket or sitting at the dinner table. There is ample evidence that situational factors influence momentary goals and preferences, and thereby food choice (102, 103). For example, in-store communication and cues at the consumption site can trigger hedonic- or health-related goals and thereby steer choices toward goal-congruent alternatives (104–108). The above-mentioned contextual cues, which are normally present at the point of purchase, are lacking in most fMRI studies. However, possibly more problematic, situational factors in the fMRI research setting, like seeing medical equipment, might activate associated information (i.e., thoughts about disease, medical treatments) and influence current goals (e.g., prevention of disease) itself and thereby influence behavior. It is unknown how the presence of medical equipment influences food choice and underlying cognition, and this is a relevant topic for further study. Further, given the strong effects of situational factors on choice and potentially on the neural processes leading to that choice, it is important that authors describe the complete study setting with a high level of detail. For example, it should minimally be mentioned whether the experiment was carried out at a hospital or at a research-dedicated MRI scanner in a nonmedical facility.

Aside from these situational factors, fMRI food choice tasks are generally highly simplified, showing (cut-out) images on a plain background, and are thus very different from the real-life food-choice environment (109–113). Situational and task-related factors combined might result in very different choices in fMRI research than in real life. If choice behavior differs between fMRI tasks and real life, how can we be confident that the cognitive process we measure during choice is the one we actually aim to measure? So far, to our knowledge, only a few studies have related choices made in the scanner to a "real-life" measure of eating behavior; namely, intake at a subsequent ad libitum laboratory buffet meal (114) and intake at a buffet lunch the next day (115). In the former study, however, in-scanner

choices were not related to intake at the buffet. To assess how representative food-choice behavior in fMRI tasks is for real-life food choices, future studies should incorporate real-life measures of eating behavior and relate these to in-scanner behaviors. This will allow us to establish the need for more realistic fMRI food-choice paradigms.

One approach to develop more realistic fMRI paradigms is by using virtual reality (VR). VR provides the ultimate level of immersion, creating a sense of physical presence in the 3D virtual environment, and has been successfully applied in a wide range of fields including psychiatry and medicine (116, 117). Moreover, in recent years, several virtual supermarkets have been developed (118–120), which enables collection of purchase data in a very controlled yet realistic environment. VR has a major potential for use in neuroimaging food-choice research because individuals quickly feel "embedded" in VR environments, such that the actual situation (lying in an MRI scanner) is suppressed in favor of the virtual situation (walking in the supermarket) (121). Several studies have shown that purchasing behavior in virtual supermarkets is relatively similar to actual purchase behavior (122–125). However, increased realism might come at the cost of increased noise and excessive visual stimulation, which might decrease sensitivity to detect signals of interest. To our knowledge, to date only one virtual supermarket paradigm that can be used in fMRI research has been developed (http://nutritionalneuroscience.eu/index.php/resources/neuroshop-virtual-supermarket). In this paradigm, participants can first freely navigate through the virtual supermarket with a joystick. This serves to embed the participant in the virtual supermarket and foster involvement in the task of grocery shopping. Subsequently, participants perform a more standardized fMRI choice task in which shelves with the same design are shown, and choice blocks are interspersed with movies of walking around from shelf to shelf, in order to maintain embedding. This provides a first step toward exploiting the potential of VR to produce more ecologically valid measures of food choice and underlying neural processes.

More realistic feeding paradigms.

To better mimic ingestive behavior, there is a need to move beyond stimulation with passive reception of small boluses of liquid. The major hurdle here has been the sensitivity of fMRI to movement. However, recent advances in hardware and software offer hope that sequences can be compiled that will be more robust and perhaps even allow us to measure responses to active sipping, swallowing, and even chewing solid foods. For example, multiecho fMRI increases the signal-to-noise ratio by a factor of 4 (126), whereas multiband acquisition provides enhanced speed to increase the temporal resolution, allowing greater ability to deconvolve the BOLD response in the context of movement. Also in development is echo planar imaging with the "keyhole technique," which increases the signal readout even further, allowing 25–30% increases in either spatial or temporal resolution. These improvements in data acquisition can then be coupled to new technology enabling delivery of solid foods to participants lying in the scanner bore. Although there is some way to go, and chewing poses an additional risk of movement artifacts as well as aliasing of activity from the temporalis muscles, such technologies are on the horizon (127).

Discussion

The potential of functional neuroimaging for leveraging our understanding of the drivers of eating behavior is substantial because it can elucidate the underlying neural processes and how these are affected by the diverse determinants of eating behavior. However, to maximize the yield of neuroimaging methods, it is of paramount importance to adhere to high standards in terms of experimental and task design, and subsequent data analysis to ensure sufficient detection power, specificity, and interpretability. To accommodate the complexity of nutrition research and to be able to distinguish noise from meaningful variability, the use of standardized methods, proper phenotyping, and reporting of sufficient methodological detail are necessary to enhance data pooling and meta-analyses of nutritional imaging data. Moreover, there is a need for more prospective and repeated-measures studies to elucidate etiology and establish neural markers so as to provide novel and specific targets for intervention.

None of the authors has a potential conflict of interest. The authors' responsibilities were as follows—All authors designed and wrote the paper; PAMS had primary responsibility for final content; and all authors read and approved the final manuscript.

References

1. Small DM, Jones-Gotman M, Zatorre RJ, Petrides M, Evans AC. Flavor processing: More than the sum of its parts. Neuroreport 1997;8(18):3913–7.
2. Cerf B, Lebihan D, Van de Moortele PF, Mac LP, Faurion A. Functional lateralization of human gustatory cortex related to handedness disclosed by fMRI study. Ann NY Acad Sci 1998;855:575–8.
3. Tataranni PA, Gautier JF, Chen K, Uecker A, Bandy D, Salbe AD, Pratley RE, Lawson M, Reiman EM, Ravussin E. Neuroanatomical correlates of hunger and satiation in humans using positron emission tomography. Proc Natl Acad Sci USA 1999;96(8):4569–74.
4. Gautier JF, Chen K, Salbe AD, Bandy D, Pratley RE, Heiman M, Ravussin E, Reiman EM, Tataranni PA. Differential brain responses to satiation in obese and lean men. Diabetes 2000;49(5):838–46.
5. van der Laan LN, de Ridder DT, Viergever MA, Smeets PA. The first taste is always with the eyes: A meta-analysis on the neural correlates of processing visual food cues. Neuroimage 2011;55(1):296–303.
6. Drew Sayer R, Tamer GG, Jr., Chen N, Tregellas JR, Cornier MA, Kareken DA, Talavage TM, McCrory MA, Campbell WW. Reproducibility assessment of brain responses to visual food stimuli in adults with overweight and obesity. Obesity (Silver Spring) 2016;24(10):2057–63.
7. Karra E, O'Daly OG, Choudhury AI, Yousseif A, Millership S, Neary MT, Scott WR, Chandarana K, Manning S, Hess ME, et al. A link between FTO, ghrelin, and impaired brain food-cue responsivity. J Clin Invest 2013;123(8):3539–51.
8. Heni M, Kullmann S, Veit R, Ketterer C, Frank S, Machicao F, Staiger H, Haring HU, Preissl H, Fritsche A. Variation in the obesity risk gene FTO determines the postprandial cerebral processing of food stimuli in the prefrontal cortex. Mol Metab 2014;3(2):109–13.
9. Kuhn AB, Feis DL, Schilbach L, Kracht L, Hess ME, Mauer J, Bruning JC, Tittgemeyer M. FTO gene variant modulates the neural correlates of visual food perception. Neuroimage 2016;128:21–31.
10. Rapuano KM, Zieselman AL, Kelley WM, Sargent JD, Heatherton TF, Gilbert-Diamond D. Genetic risk for obesity predicts nucleus accumbens size and responsivity to real-world food cues. Proc Natl Acad Sci USA 2017;114(1):160–5.
11. Wiemerslage L, Nilsson EK, Solstrand Dahlberg L, Ence-Eriksson F, Castillo S, Larsen AL, Bylund SB, Hogenkamp PS, Olivo G,

Bandstein M, et al. An obesity-associated risk allele within the FTO gene affects human brain activity for areas important for emotion, impulse control and reward in response to food images. Eur J Neurosci 2016;43(9):1173–80.
12. van der Laan LN, Smeets PAM. You are what you eat: A neuroscience perspective on consumers' personality characteristics as determinants of eating behavior. Curr Opin Food Sci 2015;3(0):11–8.
13. Carp J. The secret lives of experiments: Methods reporting in the fMRI literature. Neuroimage 2012;63(1):289–300.
14. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafo MR. Power failure: Why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013;14(5):365–76.
15. Mumford JA. A power calculation guide for fMRI studies. Soc Cogn Affect Neur 2012;7(6):738–42.
16. Durnez J, Degryse J, Moerkerke B, Seurinck R, Sochat V, Poldrack R, Nichols T. Power and sample size calculations for fMRI studies based on the prevalence of active peaks. bioRxiv 2016. doi:10.1101/049429.
17. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci USA 2016;113(28):7900–5.
18. O'Reilly JX, Woolrich MW, Behrens TEJ, Smith SM, Johansen-Berg H. Tools of the trade: Psychophysiological interactions and functional connectivity. Soc Cogn Affect Neur 2012;7(5):604–9.
19. Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. A default mode of brain function. Proc Natl Acad Sci USA 2001;98(2):676–82.
20. Buckner RL, Andrews-Hanna JR, Schacter DL. The brain's default network: Anatomy, function, and relevance to disease. Ann NY Acad Sci 2008;1124:1–38.
21. Spetter MS, de Graaf C, Viergever MA, Smeets PA. Anterior cingulate taste activation predicts ad libitum intake of sweet and savory drinks in healthy, normal-weight men. J Nutr 2012;142(4): 795–802.
22. Mehta S, Melhorn SJ, Smeraglio A, Tyagi V, Grabowski T, Schwartz MW, Schur EA. Regional brain response to visual food cues is a marker of satiety that predicts food choice. Am J Clin Nutr 2012;96(5):989–99.
23. Murdaugh DL, Cox JE, Cook EW, III, Weller RE. fMRI reactivity to high-calorie food pictures predicts short- and long-term outcome in a weight-loss program. Neuroimage 2012;59(3):2709–21.
24. Demos KE, Heatherton TF, Kelley WM. Individual differences in nucleus accumbens activity to food and sexual images predict weight gain and sexual behavior. J Neurosci 2012;32(16):5549–52.
25. Yokum S, Ng J, Stice E. Attentional bias to food images associated with elevated weight and future weight gain: An fMRI study. Obesity (Silver Spring) 2011;19(9):1775–83.
26. Yokum S, Gearhardt AN, Harris JL, Brownell KD, Stice E. Individual differences in striatum activity to food commercials predict weight gain in adolescents. Obesity (Silver Spring) 2014;22(12):2544–51.
27. Stice E, Burger KS, Yokum S. Reward region responsivity predicts future weight gain and moderating effects of the TaqIA allele. J Neurosci 2015;35(28):10316–24.
28. Volkow ND, Wang GJ, Telang F, Fowler JS, Thanos PK, Logan J, Alexoff D, Ding YS, Wong C, Ma Y, et al. Low dopamine striatal D2 receptors are associated with prefrontal metabolism in obese subjects: Possible contributing factors. Neuroimage 2008;42(4): 1537–43.
29. Wang GJ, Volkow ND, Logan J, Pappas NR, Wong CT, Zhu W, Netusil N, Fowler JS. Brain dopamine and obesity. Lancet 2001;357(9253):354–7.
30. Dang LC, Samanez-Larkin GR, Castrellon JJ, Perkins SF, Cowan RL, Zald DH. Associations between dopamine D2 receptor availability and BMI depend on age. Neuroimage 2016;138:176–83.
31. Benton D, Young HA. A meta-analysis of the relationship between brain dopamine receptors and obesity: A matter of changes in behavior rather than food addiction? Int J Obes (Lond) 2016;40 Suppl 1:S12–21.
32. Schwartz MW, Seeley RJ, Zeltser LM, Drewnowski A, Ravussin E, Redman LM, Leibel RL. Obesity pathogenesis: An Endocrine Society scientific statement. Endocr Rev 2017;38(4):267–96.
33. Braff DL. The importance of endophenotypes in schizophrenia research. Schizophr Res 2015;163(1–3):1–8.
34. Willette AA, Kapogiannis D. Does the brain shrink as the waist expands? Ageing Res Rev 2015;20:86–97.
35. Poldrack RA, Fletcher PC, Henson RN, Worsley KJ, Brett M, Nichols TE. Guidelines for reporting an fMRI study. Neuroimage 2008;40(2):409–14.
36. Nichols TE, Das S, Eickhoff SB, Evans AC, Glatard T, Hanke M, Kriegeskorte N, Milham MP, Poldrack RA, Poline JB, et al. Best practices in data analysis and sharing in neuroimaging using MRI. Nat Neurosci 2017;20(3):299–303.
37. Ioannidis JP. Why most published research findings are false. PLoS Med 2005;2(8):e124.
38. Cremers HR, Wager TD, Yarkoni T. The relation between statistical power and inference in fMRI. PLoS One 2017;12(11):e0184923.
39. Yarkoni T. Big correlations in little studies: Inflated fMRI correlations reflect low statistical power-commentary on Vul et al. (2009). Perspect Psychol Sci 2009;4(3):294–8.
40. Crutzen R, Peters GY. Targeting next generations to change the common practice of underpowered research. Front Psychol 2017;8:1184.
41. Joyce KE, Hayasaka S. Development of PowerMap: A software package for statistical power calculation in neuroimaging studies. Neuroinformatics 2012;10(4):351–65.
42. Poldrack RA, Farah MJ. Progress and challenges in probing the human brain. Nature 2015;526(7573):371–9.
43. Tong F, Pratte MS. Decoding patterns of human brain activity. Annu Rev Psychol 2012;63:483–509.
44. Bassett DS, Sporns O. Network neuroscience. Nat Neurosci 2017;20(3):353–64.
45. Calhoun VD, Miller R, Pearlson G, Adali T. The chronnectome: Time-varying connectivity networks as the next frontier in fMRI data discovery. Neuron 2014;84(2):262–74.
46. Cornier MA, Despres JP, Davis N, Grossniklaus DA, Klein S, Lamarche B, Lopez-Jimenez F, Rao G, St-Onge MP, Towfighi A, et al. Assessing adiposity: A scientific statement from the American Heart Association. Circulation 2011;124(18):1996–2019.
47. Rosenbaum M, Agurs-Collins T, Bray MS, Hall KD, Hopkins M, Laughlin M, MacLean PS, Maruvada P, Savage CR, Small DM, et al. Accumulating Data to Optimally Predict Obesity Treatment (ADOPT): Recommendations from the biological domain. Obesity (Silver Spring) 2018;26 Suppl 2:S25–34.
48. LaBar KS, Gitelman DR, Parrish TB, Kim YH, Nobre AC, Mesulam MM. Hunger selectively modulates corticolimbic activation to food stimuli in humans. Behav Neurosci 2001;115(2):493–500.
49. Goldstone AP, de Hernandez CG, Beaver JD, Muhammed K, Croese C, Bell G, Durighel G, Hughes E, Waldman AD, Frost G, et al. Fasting biases brain reward systems towards high-calorie foods. Eur J Neurosci 2009;30(8):1625–35.
50. Frank S, Laharnar N, Kullmann S, Veit R, Canova C, Hegner YL, Fritsche A, Preissl H. Processing of food pictures: Influence of hunger, gender and calorie content. Brain Res 2010;1350:159–66.
51. Charbonnier L, van Meer F, Johnstone AM, Crabtree D, Buosi W, Manios Y, Androutsos O, Giannopoulou A, Viergever MA, Smeets PAM. Effects of hunger state on the brain responses to food cues across the life span. Neuroimage 2018;171:246–55.
52. Hernandez EA, Kahl S, Seelig A, Begovatz P, Irmler M, Kupriyanova Y, Nowotny B, Nowotny P, Herder C, Barosa C, et al. Acute dietary fat intake initiates alterations in energy metabolism and insulin resistance. J Clin Invest 2017;127(2):695–708.
53. Lichtman SW, Pisarska K, Berman ER, Pestone M, Dowling H, Offenbacher E, Weisel H, Heshka S, Matthews DE, Heymsfield SB. Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. N Engl J Med 1992;327(27):1893–8.
54. Abdallah L, Chabert M, Louis-Sylvestre J. Cephalic phase responses to sweet taste. Am J Clin Nutr 1997;65(3):737–43.
55. Huettel S, Song A, McCarthy G. Functional magnetic resonance imaging. Sunderland, MA: Sinauer Associates; 2004.
56. Bennett C, Miller M, Wolford G. Neural correlates of interspecies perspective taking in the post-mortem Atlantic salmon: An argument for multiple comparisons correction. Neuroimage 2009;47:S125.
57. Woo CW, Krishnan A, Wager TD. Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. Neuroimage 2014;91:412–9.
58. Hayasaka S, Nichols TE. Validating cluster size inference: Random field and permutation methods. Neuroimage 2003;20(4):2343–56.
59. Hodgson K, Poldrack RA, Curran JE, Knowles EE, Mathias S, Göring HH, Yao N, Olvera RL, Fox PT, Almasy L. Shared genetic factors

influence head motion during MRI and body mass index. Cereb Cortex 2016:1–8.
60. Nakamura Y, Goto TK, Tokumori K, Yoshiura T, Kobayashi K, Nakamura Y, Honda H, Ninomiya Y, Yoshiura K. Localization of brain activation by umami taste in humans. Brain Res 2011;1406:18–29.
61. Small DM, Gregory MD, Mak YE, Gitelman D, Mesulam MM, Parrish T. Dissociation of neural representation of intensity and affective valuation in human gustation. Neuron 2003;39(4):701–11.
62. Friston KJ, Williams S, Howard R, Frackowiak RS, Turner R. Movement-related effects in fMRI time-series. Magn Reson Med 1996;35(3):346–55.
63. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 2012;59(3):2142–54.
64. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 2014;84:320–41.
65. Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, Petersen SE. Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. Hum Brain Mapp 2014;35(5):1981–96.
66. Martin RE, Goodyear BG, Gati JS, Menon RS. Cerebral cortical representation of automatic and volitional swallowing in humans. J Neurophysiol 2001;85(2):938–50.
67. Veldhuizen MG, Small DM. Modality-specific neural effects of selective attention to taste and odor. Chem Senses 2011;36(8):747–60.
68. Dipasquale O, Sethi A, Lagana MM, Baglio F, Baselli G, Kundu P, Harrison NA, Cercignani M. Comparing resting state fMRI de-noising approaches using multi- and single-echo acquisitions. PLoS One 2017;12(3):e0173289.
69. Brooks JCW, Faull OK, Pattinson KTS, Jenkinson M. Physiological noise in brainstem fMRI. Front Hum Neurosci 2013;7:623.
70. Pruim RHR, Mennes M, Buitelaar JK, Beckmann CF. Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. Neuroimage 2015;112:278–87.
71. Singer JDW, Willett JB. Applied longitudinal data analysis: modeling change and event occurrence. New York: Oxford University Press; 2003.
72. Geha PY, Aschenbrenner K, Felsted J, O'Malley SS, Small DM. Altered hypothalamic response to food in smokers. Am J Clin Nutr 2013;97(1):15–22.
73. Measelle JR, Stice E, Hogansen JM. Developmental trajectories of co-occurring depressive, eating, antisocial, and substance abuse problems in female adolescents. J Abnorm Psychol 2006;115(3):524–38.
74. Stice E, Yokum S. Gain in body fat is associated with increased striatal response to palatable food cues, whereas body fat stability is associated with decreased striatal response. J Neurosci 2016;36(26):6949–56.
75. Stice E, Yokum S, Veling H, Kemps E, Lawrence NS. Pilot test of a novel food response and attention training treatment for obesity: Brain imaging data suggest actions shape valuation. Behav Res Ther 2017;94:60–70.
76. Cornier MA, Salzberg AK, Endly DC, Bessesen DH, Tregellas JR. Sex-based differences in the behavioral and neuronal responses to food. Physiol Behav 2010;99(4):538–43.
77. Kishinevsky FI, Cox JE, Murdaugh DL, Stoeckel LE, Cook EW, III, Weller RE. fMRI reactivity on a delay discounting task predicts weight gain in obese women. Appetite 2012;58(2):582–92.
78. Stice E, Yokum S, Bohon C, Marti N, Smolen A. Reward circuitry responsivity to food predicts future increases in body mass: Moderating effects of DRD2 and DRD4. Neuroimage 2010;50(4):1618–25.
79. Winter SR, Yokum S, Stice E, Osipowicz K, Lowe MR. Elevated reward response to receipt of palatable food predicts future weight variability in healthy-weight adolescents. Am J Clin Nutr 2017;105(4):781–9.
80. Ness A, Bruce J, Bruce A, Aupperle R, Lepping R, Martin L, Hancock L, Patrician T, Malley S, Selim N, et al. Pre-surgical cortical activation to food pictures is associated with weight loss following bariatric surgery. Surg Obes Relat Dis 2014;10(6):1188–95.
81. Copas JB. Using regression models for prediction: Shrinkage and regression to the mean. Stat Meth Med Res 1997;6(2):167–83.
82. Varoquaux G. Cross-validation failure: Small sample sizes lead to large error bars. Neuroimage 2018;180:68–77.
83. Skocik M, Collins J, Callahan-Flintoft C, Bowman H, Wyble B. I tried a bunch of things: The dangers of unexpected overfitting in classification. bioRxiv 2016. doi:10.1101/078816.
84. Poldrack RA, Gorgolewski KJ. Making big data open: Data sharing in neuroimaging. Nat Neurosci 2014;17(11):1510–7.
85. Fox PT, Lancaster JL, Laird AR, Eickhoff SB. Meta-analysis in human neuroimaging: Computational modeling of large-scale databases. Annu Rev Neurosci 2014;37:409–34.
86. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 2011;8(8):665–70.
87. Salimi-Khorshidi G, Smith SM, Keltner JR, Wager TD, Nichols TE. Meta-analysis of neuroimaging data: A comparison of image-based and coordinate-based pooling of studies. Neuroimage 2009;45(3):810–23.
88. Gorgolewski KJ, Varoquaux G, Rivera G, Schwarz Y, Ghosh SS, Maumet C, Sochat VV, Nichols TE, Poldrack RA, Poline J-B, et al. NeuroVault.org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain. Front Neuroinform 2015;9(8). doi:10.3389/fninf.2015.00008.
89. Van Essen DC, Smith SM, Barch DM, Behrens TE, Yacoub E, Ugurbil K. The WU-Minn Human Connectome Project: An overview. Neuroimage 2013;80:62–79.
90. Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME, Toro R, Jahanshad N, Schumann G, Franke B, et al. The ENIGMA Consortium: Large-scale collaborative analyses of neuroimaging and genetic data. Brain Imaging Behav 2014;8(2):153–82.
91. Para AE, Sam K, Poublanc J, Fisher JA, Crawley AP, Mikulis DJ. Invalidation of fMRI experiments secondary to neurovascular uncoupling in patients with cerebrovascular disease. J Magn Reson Imaging: JMRI 2017;46(5):1448–55.
92. Hege MA, Veit R, Krumsiek J, Kullmann S, Heni M, Rogers PJ, Brunstrom JM, Fritsche A, Preissl H. Eating less or more—Mindset induced changes in neural correlates of pre-meal planning. Appetite 2018;125:492–501.
93. van Rijn I, de Graaf C, Smeets PA. It's in the eye of the beholder: Selective attention to drink properties during tasting influences brain activation in gustatory and reward regions. Brain Imaging Behav 2018;12(2):425–36.
94. Veldhuizen MG, Small DM. Modality-specific neural effects of selective attention to taste and odor. Chem Senses 2011;36(8):747–60.
95. Grabenhorst F, Rolls ET. Selective attention to affective value alters how the brain processes taste stimuli. Eur J Neurosci 2008;27(3):723–9.
96. Poldrack RA. Can cognitive processes be inferred from neuroimaging data? Trends Cogn Sci 2006;10(2):59–63.
97. Zhu JN, Wang JJ. The cerebellum in feeding control: Possible function and mechanism. Cell Mol Neurobiol 2008;28(4):469–78.
98. Frank S, Kullmann S, Veit R. Food related processes in the insular cortex. Front Hum Neurosci 2013;7:499.
99. Reid AT, Bzdok D, Genon S, Langner R, Muller VI, Eickhoff CR, Hoffstaedter F, Cieslik EC, Fox PT, Laird AR, et al. ANIMA: A data-sharing initiative for neuroimaging meta-analyses. Neuroimage 2016;124(Pt B):1245–53.
100. Smeets PA, Charbonnier L, van Meer F, van der Laan LN, Spetter MS. Food-induced brain responses and eating behaviour. Proc Nutr Soc 2012;71(4):511–20.
101. Gorgolewski KJ, Alfaro-Almagro F, Auer T, Bellec P, Capota M, Chakravarty MM, Churchill NW, Cohen AL, Craddock RC, Devenyi GA, et al. BIDS apps: Improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. PLoS Comput Biol 2017;13(3):e1005209.
102. Papies EK. Situating interventions to bridge the intention–behaviour gap: A framework for recruiting nonconscious processes for behaviour change. Social Pers Psychol Compass 2017;11(7):e12323.
103. Stroebele N, De Castro JM. Effect of ambience on food intake and food choice. Nutrition 2004;20(9):821–38.
104. van der Laan LN, Papies EK, Hooge IT, Smeets PA. Goal-directed visual attention drives health goal priming: An eye-tracking experiment. Health Psychol 2017;36(1):82–90.

105. Papies EK, Hamstra P. Goal priming and eating behavior: Enhancing self-regulation by environmental cues. Health Psychol 2010;29(4):384–8.
106. Papies EK. Health goal priming as a situated intervention tool: How to benefit from nonconscious motivational routes to health behaviour. Health Psychol Rev 2016;10(4):408–24.
107. Harris JL, Bargh JA, Brownell KD. Priming effects of television food advertising on eating behavior. Health Psychol 2009;28(4):404–13.
108. Ouwehand C, Papies EK. Eat it or beat it. The differential effects of food temptations on overweight and normal-weight restrained eaters. Appetite 2010;55(1):56–60.
109. van der Laan LN, de Ridder DT, Viergever MA, Smeets PA. Activation in inhibitory brain regions during food choice correlates with temptation strength and self-regulatory success in weight-concerned women. Front Neurosci 2014;8:308.
110. van der Laan LN, de Ridder DT, Charbonnier L, Viergever MA, Smeets PA. Sweet lies: Neural, visual, and behavioral measures reveal a lack of self-control conflict during food choice in weight-concerned women. Front Behav Neurosci 2014;8:184.
111. Charbonnier L, van der Laan LN, Viergever MA, Smeets PAM. Functional MRI of challenging food choices: Forced choice between equally liked high- and low-calorie foods in the absence of hunger. PLoS One 2015;10(7):e0131727.
112. van der Laan LN, de Ridder DT, Viergever MA, Smeets PA. Appearance matters: Neural correlates of food choice and packaging aesthetics. PLoS One 2012;7(7):e41738.
113. Hare TA, Camerer CF, Rangel A. Self-control in decision-making involves modulation of the vmPFC valuation system. Science 2009;324(5927):646–8.
114. Medic N, Ziauddeen H, Forwood SE, Davies KM, Ahern AL, Jebb SA, Marteau TM, Fletcher PC. The presence of real food usurps hypothetical health value judgment in overweight people. eNeuro 2016;3(2). doi:10.1523/eneuro.0025-16.2016.
115. Foerde K, Steinglass JE, Shohamy D, Walsh BT. Neural mechanisms supporting maladaptive food choices in anorexia nervosa. Nat Neurosci 2015;18(11):1571–3.
116. Schwebel DC, McClure LA, Severson J. Usability and feasibility of an internet-based virtual pedestrian environment to teach children to cross streets safely. Virtual Real 2014;18(1):5–11.
117. Gasco J, Holbrook TJ, Patel A, Smith A, Paulson D, Muns A, Desai S, Moisi M, Kuo YF, Macdonald B, et al. Neurosurgery simulation in residency training: Feasibility, cost, and educational benefit. Neurosurgery 2013;73 Suppl 1:39–45.
118. Nederkoorn C, Guerrieri R, Havermans RC, Roefs A, Jansen A. The interactive effect of hunger and impulsivity on food intake and purchase in a virtual supermarket. Int J Obes (Lond) (2005) 2009;33(8):905–12.
119. Riva G, Gaggioli A, Villani D, Preziosa A, Morganti F, Corsi R, Faletti G, Vezzadini L. NeuroVR: An open source virtual reality platform for clinical psychology and behavioral neurosciences. Stud Health Technol Inform 2007;125:394–9.
120. Waterlander WE, Scarpa M, Lentz D, Steenhuis IH. The virtual supermarket: An innovative research tool to study consumer food purchasing behaviour. BMC Public Health 2011;11:589.
121. Lenggenhager B, Tadi T, Metzinger T, Blanke O. Video ergo sum: Manipulating bodily self-consciousness. Science 2007;317(5841):1096–9.
122. Burke RR, Harlam BA, Kahn BE, Lodish LM. Comparing dynamic consumer choice in real and computer-simulated environments. J Consum Res 1992;19(1):71–82.
123. van Herpen E, van den Broek E, van Trijp HCM, Yu T. Can a virtual supermarket bring realism into the lab? Comparing shopping behavior using virtual and pictorial store representations to behavior in a physical store. Appetite 2016;107:196–207.
124. Waterlander WE, Jiang Y, Steenhuis IH, Ni Mhurchu C. Using a 3D virtual supermarket to measure food purchase behavior: A validation study. J Med Internet Res 2015;17(4):e107.
125. Campo KG, Gijsbrechts E, Guerra F. Computer simulated shopping experiments for analyzing dynamic purchasing patterns: Validation and guidelines. J Empir Gen Mark Sci 1999;4(2):22–61.
126. Kundu P, Voon V, Balchandani P, Lombardo MV, Poser BA, Bandettini PA. Multi-echo fMRI: A review of applications in fMRI denoising and analysis of BOLD signals. Neuroimage 2017;154:59–80.
127. Yun SD, Shah NJ. Whole-brain high in-plane resolution fMRI using accelerated EPIK for enhanced characterisation of functional areas at 3T. PLoS One 2017;12(9):e0184759.