Human brain responses to gustatory and food stimuli: A meta-evaluation of neuroimaging meta-analyses

Andy Wai Kan Yeung[a,*], Natalie Sui Miu Wong[b], Hakwan Lau[c,d,e,f], Simon B. Eickhoff[g,h]

[a] Oral and Maxillofacial Radiology, Applied Oral Sciences and Community Dental Care, Faculty of Dentistry, The University of Hong Kong, Hong Kong, China (ndyeung@hku.hk)

[b] Oral and Maxillofacial Surgery, Faculty of Dentistry, The University of Hong Kong, Hong Kong, China (smwong26@hku.hk)

[c] Department of Psychology, UCLA, Los Angeles, USA

[d] Brain Research Institute, UCLA, Los Angeles, USA

[e] Department of Psychology, The University of Hong Kong, Hong Kong, China

[f] State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, China

[g] Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany (Simon.Eickhoff@uni-duesseldorf.de)

[h] Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Centre Jülich, Jülich, Germany

*Corresponding author:

Andy Wai Kan Yeung, 1B39B, Prince Philip Dental Hospital, 34 Hospital Road, Hong Kong.

Phone: 852-2859-0403

Email: ndyeung@hku.hk

Highlights

- Neuroimaging meta-analyses on gustation / food / taste stimuli were evaluated.
- All of them were of moderate and high quality of evidence.
- Meta-analyses are increasingly adopting more stringent statistical thresholds.
- Newer meta-analyses tended to report using standard MNI space coordinates.
- The correct implementation of GingerALE software should be ensured.

Abstract

Multiple neuroimaging meta-analyses have been published concerning gustation, food and taste. A meta-evaluation of these meta-analyses was conducted to qualitatively evaluate the presented evidence. A systematic search was done using multiple databases, in which no restriction was placed on participants and nature of interventions (stimuli vs control). Twenty-three meta-analyses were identified and analyzed. All of them have met 4–9 criteria, out of 11, from the modified checklist constructed by Müller et al. (2018), which implied moderate to high quality of evidence. One of the concerns we found was that no meta-analysis surveyed had been explicitly pre-registered. Also, only three meta-analyses (13.0%) provided clear explanation of how they accounted for sample overlap. Only six meta-analyses (26.1%) explicitly described how they double checked the data. Only two of the 20 meta-analyses (10.0%) using GingerALE software used both the debugged version (v2.3.6) as well as the recommended cluster-level inference with familywise error rate correction. Overall, meta-analyses are increasingly adopting more stringent statistical thresholds, but unfortunately not larger number of studies contained in the analyses.

Keywords: meta-evaluation; meta-analysis; gustation; food; taste; neuroimaging; fMRI; literature analysis; activation likelihood estimation; review

1. Introduction

The ability to taste and eat is essential for our daily life, and accordingly heavily investigated. Techniques such as functional magnetic resonance imaging (fMRI) have enabled scientists to measure the neural responses induced by food stimuli as presented in various forms, such as visual cues (Rothemund et al., 2007), pure tastants (Small et al., 2003), real food (Stice et al., 2008), and mental imagery (Barrós-Loscertales et al., 2011). Moreover, experiments have reported altered brain activation in different subject groups, such as patients with anorexia nervosa (Santel et al., 2006), bulimia nervosa (Brooks et al., 2011), and obesity (Ng et al., 2011). With such a rich and diversified literature in gustation, numerous neuroimaging meta-analyses have been published in an attempt to summarize the existing findings. All these meta-analyses provided important findings from scientists to advance our knowledge and understanding towards the neurobiology of tasting and eating behavior in healthy and patient populations. For instance, earlier meta-analyses on taste neuroimaging studies confirmed that the insula is functioning as the primary taste cortex (Veldhuizen et al., 2011; Yeung et al., 2017). Further, it was found that the left anterior insula is the only brain region that consistently activates in response to all taste, odor, and relevant images, three different kinds of food cues (Huerta et al., 2014). A recent meta-analysis further reported that the insula is actually responsible for processing different aspects of taste, including quality, intensity, and pleasantness (Yeung et al., 2018). From these examples we can see that meta-analyses have been instrumental in helping us form a general and informative picture of the neurobiology of tasting.

In this fast-growing literature, one outstanding concern is that the heterogeneous stimuli and populations employed in neuroimaging studies might potentially lead to different findings. Clinicians, scientists and taste researchers often need to compare and contrast all these

findings to assess how systematically the key results may change accordingly, when stimuli and populations change. This is important both for disease diagnosis as well as food product improvement. This further highlights other critical contributions by meta-analyses.

The conclusions of a meta-analysis may be strongly influenced by the quality of the meta-analytic process (Nakagawa et al., 2017). Because of that, reporting guidelines were developed to facilitate the synthesis of meta-analyses, perhaps with the most notable examples being the Quality of Reporting of Meta-analyses (QUOROM) developed in the late 1990s (Moher et al., 2000), and its successor published in 2009, the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (Moher et al., 2009). Both guidelines apply to meta-analyses based on human studies, with the latter composes of a 27-item checklist and a flow diagram, which help authors synthesize meta-analytic reports with the essential details included and reported in a standardized manner. The items concern about the reporting details of a meta-analysis' title, abstract, introduction, methods, results, discussion, and funding. Compared to QUOROM, PRISMA addresses more details, such as if a meta-analysis has a protocol and how it can be accessed, how the risk of bias of the included studies is assessed, and the information on the sources of funding (Moher et al., 2009). Meanwhile, another review paper was recently published to assist researchers to assess the quality of meta-analyses based on data from non-human species (Nakagawa et al., 2017). In various fields of biological and health sciences, studies have already been published to evaluate the quality of meta-analyses in the literature according to these guidelines, including gastroenterology and hepatology (Panic et al., 2013), ear, nose and throat (Peters et al., 2015), and dentistry (Bijle et al., 2018). These examples illustrated that there are "evidence classes" for meta-analyses in the form of a score or percentage that reflects the adherence to a particular guideline, allowing readers to better comprehend the limitations of

the meta-analytic results that are naturally liable to various biases and the problem of "garbage in, garbage out" (Egger et al., 2001).

In neuroimaging, however, we used to have a lack of specific guidelines for performing meta-analyses, and we tended to see every meta-analysis as equal in terms of scientific quality. The transparency, traceability, replicability and reporting of the meta-analytic results were not evaluated. Therefore, in the current meta-evaluation, we aimed to identify and evaluate the gustatory neuroimaging meta-analyses according to the recommendations of a recently published paper (Müller et al., 2018), which is a consensus guideline by all major developers that sets the new standard for neuroimaging (coordinate-based) meta-analyses. In merely more than a year since its publication, Müller et al. (2018) has already accumulated 30 citations according to Google Scholar, with many latest neuroimaging meta-analyses citing and adhering to the guideline. Using this meta-evaluation on taste meta-analyses, we attempted to investigate current research trends in neuroimaging studies on the topic. We attempted to investigate how strictly the existing gustatory neuroimaging meta-analyses followed the guidelines set by Müller et al. (2018), their choices of brain space and statistical thresholds, and other details of their meta-analyses, such as study population, number of studies included, and types of gustatory stimuli involved.

2. Materials and methods

2.1. Protocol registration

A protocol listing the search strategy and methods of analyses was registered before the commencement of the meta-evaluation, with the PROSPERO database (registration number: CRD42018111520). There was no deviation from the specified protocol.

2.2. Database and search strategy

Four online databases were selected: Scopus, Web of Science, PsycINFO and PubMed. The search was conducted on 3 October 2018, with the search strategy: (meta-analy*) AND (food OR taste OR gustat*) AND ("brain activation" OR neuroimaging OR image-based OR coordinate-based OR "activation likelihood estimation" OR ALE OR "kernel density analysis" OR KDA OR MKDA OR "gaussian process regression" OR GPR OR "parametric voxel based meta-analysis" OR PVM OR "signed differential mapping" OR SDM OR "effect size SDM" OR ES-SDM). For Scopus and Web of Science, the title, abstract and keywords of the publication records were searched. For PsycINFO and PubMed, no search restrictions were placed. Moreover, we did not place any other restrictions, such as publication year, or language. The references of the selected meta-analyses were hand searched to identify any potentially missed meta-analyses, but no additional meta-analysis was identified during this procedure.

2.3. Inclusion and exclusion criteria

All meta-analyses reporting brain activation related to gustatory, food or taste stimuli were included in the meta-evaluation process. The PICO (participant, intervention, control, outcome) framework was considered. The participants were of any age or gender. The interventions (stimuli) could be in any form, ranging from visual (food pictures), gustatory (pure tastants or real food), to mental imagery. The control / baseline could be in any form. The outcomes were brain activations reported from the meta-analyses. A searched publication was excluded if it (1) was not related to gustation, food or taste; (2) contained no meta-analysis; or (3) reported no meta-analytic results concerning gustation, food or taste.

2.4. Meta-analysis selection process

The identification and selection process of relevant meta-analyses were carried out with an adherence to the PRISMA guideline (Figure 1). Two authors (AWKY and NSMW) independently performed the selection process, and any discrepancies were resolved through discussion and mutual consensus.

2.5. Data extraction

One author (AWKY) extracted all data to be double checked by another author (NSMW). For each meta-analysis paper, the following descriptive data were extracted: publication year, participants' mean age and body mass index (BMI), types of stimuli, outcome evaluations (within group or between group), numbers of meta-analyses performed, null results reported, original papers analyzed, foci reported, statistical threshold used, brain space used to report results, location of the results presented in the paper, relevance to researchers (if coordinates were listed, or downloadable brain maps in NIFTI format were available), and the labels of brain regions given to the reported coordinates.

2.6. Quality evaluation

The quality evaluation of the meta-analyses was conducted using the 10-item checklist proposed recently (Müller et al., 2018). Since the 9th item in the original checklist contained two main components (protocol pre-registration, and usage of default methods and parameters of established software), we separated these two components and thus made it an 11-item evaluation in the current manuscript. These items were: (1) research question answered by meta-analyses, (2) databases searched, (3) presence of inclusion and exclusion criteria, (4) accounting for sample overlap from individual studies, (5) whether the meta-analyses were based on data reported by whole-brain analyses only, (6) conversion of reported brain coordinates into a common reference space, (7) double checking of data, (8)

the presence of descriptions of included studies, (9) pre-registration of study protocol, (10) default methods and parameters of established software, and (11) types of diagnostics.

3. Results

3.1. Meta-analyses selection process

Figure 1 illustrated the workflow of the selection process. In summary, 149 publication records were identified initially through four databases, in which 74 of them were unique publications. After exclusion (see Methods), 23 meta-analysis papers were included in the current meta-evaluation.

[Insert Figure 1]

3.2. Descriptive data of the meta-analyses

Table 1 shows the descriptive data of the 23 analyzed meta-analyses. The oldest was published in 2006 and the latest in 2018. They have covered participants from the whole spectrum of BMI, from underweight (<18.5), normal (18.5–24.9), overweight (25.0–29.9) to obesity (>30.0). More of half of them (14) have involved studies with visual (14) or tastant (13) stimuli, whereas 11 of them involved real food stimuli and four involved mental imagery. The smallest meta-analysis analyzed five papers, whereas the largest one analyzed 34. All papers except one (Verhagen and Engelen, 2006) have listed the brain coordinates from meta-analyses in tabular format, and three papers (Devoto et al., 2018; Han et al., 2018; Yeung et al., 2018) have provided the resultant brain maps in NIFTI format via a listed web link or directly as supplementary materials. Incidentally, we have identified the brain maps from a fourth paper (van Meer et al., 2015), downloadable from Neurovault online depository (https://neurovault.org/collections/4211/).

198

199 [Insert Table 1]

200

201 We attempted to observe if there were any trends in the literature in reporting meta-analyses.
202 Figure 2A showed that meta-analyses in recent years exhibited increased preference to report
203 results in Montreal Neurological Institute (MNI) (Evans et al., 1993) space, over Talairach
204 and Tournoux (TAL) (Talairach and Tournoux, 1988) space. In total, 15 papers have chosen
205 MNI space, 7 have chosen TAL space, and one did not specify either. Since 2015, the ratio of
206 MNI papers to TAL papers has been 8:1. Besides, there were only three papers that
207 conducted meta-analyses with a cluster-level inference with familywise error rate (FWE)
208 correction, and coincidentally all of them were published since 2015 (Figure 2B).

209

210 [Insert Figure 2]

211

212 Meanwhile, we could not observe a trend for the number of original papers included into the
213 meta-analyses, or the number of foci reported (Figure 3). However, it seemed that meta-
214 analytic papers published in recent years tended to report a higher number of analyses, which
215 may in turn imply an increased complexity in the research questions answered. The bigger
216 studies, however, did not seem to associate with a more stringent statistical threshold, i.e. the
217 use of FWE over false-discovery rate (FDR) or uncorrected statistics.

218

219 [Insert Figure 3]

220

221 3.3. Quality evaluation of the meta-analyses

Table 2 shows the details of the evaluative outcome of the meta-analyses. The 23 analyzed meta-analyses fulfilled 4–9 of the 11 items advocated by the modified checklist from Müller et al. (2018). All studies involved coordinate-based meta-analyses except Verhagen and Engelen (2006), which only plotted all coordinates extracted from included studies onto a standard brain template "ch2" in MNI space. Meta-analyses published more recently performed similarly as those published in earlier days. The key points from Table 2 were summarized as follows:

(1) Twenty meta-analyses (87.0%) used the GingerALE software to conduct the meta-analyses by the activation likelihood estimation (ALE) method. Sixteen of them were using versions older than v2.3.6.

(2) Nineteen meta-analyses (82.6%) converted brain coordinates into a common reference space, mostly using, or presumably using, the Lancaster transform (Lancaster et al., 2007).

(3) Eighteen meta-analyses (78.3%) included data reported from whole-brain analyses only.

(4) Thirteen meta-analyses (56.5%) searched multiple databases, with PubMed being the most popular.

(5) Seven meta-analyses (30.4%) provided diagnostics of the meta-analyses, mostly by showing the contributions of experiments to significant clusters.

(6) Six meta-analyses (26.1%) explicitly described how they double checked the data.

(7) Three meta-analyses (13.0%) explained clearly on how they accounted for sample overlap.

(8) Two meta-analyses (8.7%) reported null results.

(9) No meta-analysis has explicitly stated that its study protocol had been pre-registered.

[Insert Table 2]

4. Discussion

This meta-evaluation of the neuroimaging meta-analyses concerning gustation, food and taste has revealed that all of them have fulfilled 4–9 merits, as assessed by the modified checklist recently constructed (Müller et al., 2018). According to the scoring scale (0–11) of AMSTAR assessment for systematic reviews (Pollock et al., 2017), all the 23 meta-analyses analyzed in the current study had moderate to high quality of evidence. The ALE approach was clearly the most popular meta-analytic method. It should be pointed out that the input needed for ALE is different from SDM. While ALE requires the sample size and the peak coordinates only, the minimal input for SDM is peak height, with full brain maps being more informative. It was reassuring that all analyzed studies had well-defined inclusion and exclusion criteria, and almost all of them listed the descriptive information of the analyzed studies in tabular format. These practices have enabled readers to quickly assess and determine if the research questions answered by the particular meta-analyses were relevant to them. Through the current analysis, we have noticed several issues worth of discussion, which concern the conductance and reporting of neuroimaging meta-analyses. These are elaborated as follows.

4.1. Implementation errors of dated GingerALE software

The implementation of GingerALE is an issue of some concern. The developing team of GingerALE published a report in 2017, which announced implementation errors identified in the GingerALE software prior to the version v2.3.6 (Eickhoff et al., 2017). The errors in the software code could lead to increased false positive rate. Indeed, researchers have reported that re-analyses of their original meta-analyses using the rectified version v2.3.6 have led to reduced number of activation peaks / clusters survived under the same statistical thresholds, regardless of whether the thresholds were FWE or FDR corrected (Garrison et al., 2017). Besides implementation errors, Eickhoff et al. has recently published a simulation study,

which concluded that uncorrected inference and FDR correction should be considered as invalid for ALE meta-analyses (Eickhoff et al., 2016). However, only two of the 20 evaluated papers using GingerALE software have fulfilled both criteria, i.e. the use of version v2.3.6 with cluster-level FWE inference (16 has used version prior to v2.3.6, and 17 has used FDR correction). These issues may have played a role in the nearly absence of null results reported in the 23 papers.

4.2. Pre-registration of meta-analytic protocols

The current meta-evaluation also highlights certain areas in which future studies can make improvements. For instance, authors should consider pre-registering their meta-analytic protocols (Müller et al., 2018). This is advocated because neuroimaging and neuroscience studies often involve complex statistical analytic procedures that are overly "flexible" (Carp, 2012a, b; Simmons et al., 2011). It was reported that data from an fMRI experiment could be processed by nearly 7,000 unique analytic pipelines (Carp, 2012a). Any intentional or unintentional small deviations from the pre-defined procedures and default parameters may influence the outcome, and possibly resulting in some form of "p-hacking" – an attempt to reduce the p value to make the results significant (Poldrack et al., 2017; Reddan et al., 2017) – or its unintentional equivalence. Overall, pre-registering meta-analyses at online databases (e.g. PROSPERO), to be pre-approved by journal editors (Button et al., 2013; Chambers, 2013; Chambers et al., 2017; Yeung, 2017), are recommended procedures in line with the Open Science movement endorsed by many including the current authors.

4.3. Statistical thresholding

The incorrect application of correction methodologies should be reiterated (Figure 4). Seventeen of the analyzed meta-analyses used voxel-wise FDR correction for statistical

thresholding, while computing the data with GingerALE. A decade ago, Chumbley and Friston (2009) demonstrated that FDR does not control false positive clusters well. This was reiterated in a computation study published in 2016, which demonstrated voxel-wise FDR correction in ALE meta-analyses has two shortcomings: (1) it has low sensitivity and high susceptibility to false positive results; and (2) the false positive rate at a specific brain region is strongly influenced by the strength of true convergence in other parts of the brain (Eickhoff et al., 2016). These shortcomings are detrimental to the topographic inference, which is the essence of ALE meta-analyses. Since the publication of this computation study, only two of the six subsequent meta-analyses continued to use voxel-wise FDR correction, seemingly suggesting a reduction in this choice of thresholding (88.2% down to 33.3%). Adhering to the latest recommendations (Eickhoff et al., 2016; Müller et al., 2018), we re-emphasized the use of cluster-level FWE correction of $p < 0.05$ with a cluster forming threshold of $p < 0.001$, and avoiding FDR correction, for ALE meta-analyses. Alternatively, readers should be noted that an uncorrected threshold of $p = 0.005$, with cluster extent of 10 voxels and SDM-Z > 1 is recommended for SDM meta-analyses (Müller et al., 2018; Radua et al., 2012), another common neuroimaging meta-analytic approach that seemed to be less popular in taste and food literature. It should be noted that this threshold, being uncorrected for multiple comparisons, provides only an approximation to the corrected results and can be too liberal or too conservative (Müller et al., 2018).

4.4. Inclusion of region-of-interest (ROI)-based data into meta-analyses

Four meta-analyses included ROI-based data, with one of them explicitly mentioned that the exclusion of ROI-based data did not alter the results. Two meta-analyses did not explicitly state if ROI-based data was included or not. They did not seem to report more foci than those including whole-brain data only. However, a previous meta-analysis on specific phobias

demonstrated that the inclusion of ROI-based data inflated the results hugely (Gentili et al., 2019). For instance, by comparing patients with specific phobias to healthy controls, meta-analyzing whole-brain data resulted in a significant cluster with 760 mm^3 in the anterior cingulate only; but meta-analyzing whole-brain with ROI-based data inflated the results into four clusters with a total volume of 21,912 mm^3 that covered the frontal lobe, limbic lobe, and the basal ganglia (Gentili et al., 2019). In line with this illustrative example, and the fact that the inclusion of ROI-based data violates the statistical assumptions of the meta-analytic methods, we believe that future neuroimaging meta-analyses should include data from whole-brain analyses only.

4.5. Literature search with multiple databases

In this study we found that PubMed was the most popular database used by the evaluated meta-analyses. While the PRISMA guideline has acknowledged that PubMed / MEDLINE is one of the most comprehensive sources for searching healthcare publications for conducting systematic reviews or meta-analyses, the guideline has also recommended the use of multiple databases to minimize the chance of missed studies (Liberati et al., 2009). This was echoed by one study, which reported searching by MEDLINE alone would have missed 17% of publications that met the inclusion criteria (Stevinson and Lawlor, 2004). For the 23 analyzed meta-analyses of the current manuscript, the usage of single versus multiple databases did not have an apparent effect on the number of studies included. For instance, Pursey et al. (2014) searched 9 databases, and yielded five studies; whereas Devoto et al. (2018) searched one database, and included 22 studies. As expected, the search string, research question, the inclusion / exclusion criteria and other factors could have influenced the number of records identified. Still, database choice might be one relevant factor also to be considered in future meta-analyses.

4.6. Sample overlap in multiple contrasts from a single original study

Gustatory neuroimaging studies often deliver multiple stimuli to the participants. For instance, Haase et al. (2009) recruited 18 participants and delivered six pure tastants to each of them, namely sucrose (sweet), saccharin (sweet), caffeine (bitter), citric acid (sour), guanosine 5'-monophosphate (umami), and sodium chloride (salty). The authors reported brain activations to each of these taste stimuli relative to water during hunger and satiety respectively, and to each of these taste during hunger relative to satiety. In other words, results from as many as 18 statistical contrasts were reported (6 "tastant > water" during hunger, 6 "tastant > water" during satiety, and 6 "hunger > satiety"). When a study involving this level of analytical complexity was included in a meta-analysis, should all 18 contrasts be considered as 18 independent studies with 18 participants each? This is a good example of an occasion when the issue of sample overlap should be accounted for. Müller et al. (2018) summarized the two major approaches to overcome this problem. One approach is to adjust for within-group effects, for example, by pooling brain coordinates resulted from every tastant together and treating it as a single experiment (Turkeltaub et al., 2012). An alternative option is to use data from the most representative contrast per participant group only (Cieslik et al., 2015). However, we need to also consider the purpose of the meta-analysis as the contrasts are selected. For instance, if an analysis will be done to compare the brain activations induced by natural sugar and artificial sweeteners respectively, then the coordinates resulted from "sucrose > water" should not be merged together with those from "saccharin > water". It is indeed possible for readers to deduce how the authors have accounted for the sample overlap, by counting the numbers of "experiments (or contrasts)" and "foci (or brain coordinates)" listed for each analyzed study and compared them with the originals. However, the "single group, multiple contrasts" phenomenon is quite common

among gustatory neuroimaging studies. In the current meta-evaluation, only three meta-analyses have described in clear terms how they accounted for sample overlap. Without the raw data, it is virtually impossible to assess the impact of sample overlap on the final meta-analytic results. A computational study estimated that the impact of sample overlap could affect the ALE values up to 7–9% (Turkeltaub et al., 2012). More explicit reporting of the strategies employed to overcome sample overlap in the meta-analyses, such as pooling data from multiple contrasts to treat it as one, or using data from the most representative contrast only, will be beneficial.

4.7. Conversion of coordinates into a common reference space

A different issue that needs to be addressed in meta-analyses is that, to enable meaningful pooling of data, coordinates for brain locations of activations need to be converted into a common reference space (e.g. TAL or MNI space). To ensure the meta-analytic results are replicable, it is advisable to state explicitly the algorithm used for the conversion, be it the Brett (Brett et al., 2001) or the Lancaster (Lancaster et al., 2007) transformation. In the current meta-evaluation, eight meta-analyses stated that the conversion was performed using the GingerALE software. We assume they used the recent default approach of the software, the Lancaster transformation; though the Brett transformation could have been chosen from the same droplist in the software. In the future, the provision of this kind of detail can make the meta-analytic studies more transparent and reproducible.

4.8. Study diagnostics

The provision of diagnostics may give additional detail and insight into the results. For instance, by showing contributions of experiments to each significant cluster reported, one can be more affirmative that the cluster was not resulted from a single dominating study. An

example by Yeung et al. (2017) has shown that the significant clusters were contributed by 12–38% of all experiments included; and while the sweet taste has contributed to every single cluster found in the insula, umami taste has contributions to three of them only. There seems to be no clear guideline on how contributions of experiments should be used or interpreted. Van der Laan et al. (2011) originally considered only clusters with > 50% contributing experiments. However, no cluster fulfilled the criterion. In the end, they reported all clusters, and focused the discussion on those > 33% only. Van Meer et al. (2015) similarly reported clusters with > 33% contributing experiments only, whereas Devoto et al. (2018) reported and discussed only clusters with > 3 contributing experiments. It remains an open issue for how someone should proceed when few studies drive the results.

The robustness of results can also be evaluated by jackknife analyses, which repeat the analyses again and again, with each time leaving one experiment out (Radua et al., 2012). If the type of meta-analysis takes consideration of effect sizes, then a funnel plot may be created to evaluate if the findings are largely influenced by small studies, which in turn provides insight into publication bias or robustness against publication bias (Müller et al., 2018). The use of diagnostics should allow readers to better assess if there is a high or low chance that the meta-analytic results are driven by specific studies, or by many small studies. The default analytic tools in the common meta-analytic programs were designed to address these issues of result robustness and publication bias, but not all options are available in every meta-analytic software. As mentioned in the first paragraph of the Discussion section, the inputs needed for ALE are sample size and peak coordinates, therefore diagnostics that include effect sizes are not possible for ALE. ALE allows the evaluation of contributions of experiments only, whereas jackknife analyses and funnel plot are the features of SDM. In

principle, jackknife or subsampling strategies would be possible in ALE but are not standard implemented.

In general, there seems to be a lack of literature investigating methods that modify or deviate the default parameters of the whole meta-analytic procedure including those for diagnostics. One earlier paper compared the coordinate-based meta-analytic results for various smoothing parameters, reported that the ALE was superior to other coordinate-based methods, and recommended a Gaussian kernel of 15 mm (Salimi-Khorshidi et al., 2009). However, the paper noted that the optimal setting varies from dataset to dataset (Salimi-Khorshidi et al., 2009). This is exactly the reason why GingerALE automatically derives the Gaussian kernel from the subject size inputted by the user, and has very few free parameters, to avoid users to manipulate the results (Eickhoff et al., 2009; Turkeltaub et al., 2012). This is consistent to the common notion that the flexibility of neuroimaging data analytics can introduce huge methods-related variations in the results (Carp, 2012a). However, limiting choices may lead to less insight in what a procedure exactly does. Therefore, it will be more informative if the resultant meta-analytic brain maps and original data files can be released to the public for further inspection and usage.

The aim of a neuroimaging meta-analyses is to consolidate the literature. And (maybe unfortunately), small sample studies may produce false positives contributing to that literature. Therefore, one strategy would be to include any study using any threshold, to see what findings converge. If different thresholding is used in different studies, the information of thresholding will be lost in the ALE meta-analytic procedure, but may still be assessed by SDM as the peak height information indirectly reflects the thresholding. If one really wishes to assess the effect of a particular small sample study on the ALE meta-analytic results,

perhaps one can re-run the procedures and see how the results differ from the original with the exclusion of that study.

4.9. Other recommendations

There are some miscellaneous points to be considered. One issue is that studies should explicitly report which procedures of the meta-analysis were performed independently by multiple authors, or performed multiple times by a single author (Müller et al., 2018). The former was recommended by the PRISMA guideline (Liberati et al., 2009). Meanwhile, authors of neuroimaging meta-analyses may consider uploading the resultant brain maps in standard NIFTI to the journal website as supplementary materials, or to an online depository. These brain maps would be invaluable to fellow researchers for replication, and as masks for setting ROI analyses. We also recommend that original studies should upload their brain maps to open online depositories, such as Neurovault (https://neurovault.org/). The benefit of uploading the brain maps is that the maps are important data for conducting SDM (now called Seed-based d Mapping) meta-analyses. With more brain maps available, we believe that the SDM method, which calculates the effect size, should become more popular. The shift towards image-based meta-analyses is also envisioned by the neuropsychiatric community (Tahmasian et al., 2018).

Finally, there are some practical advices that are specific for conducting gustatory meta-analyses. For instance, it is recommended to report the descriptive data of age and BMI of the participants recruited in the original studies. Only 9 and 10 out of the 23 meta-analyses reported age and BMI data respectively. Compared to older adults (mean age of 65), young adults (mean age of 23) were found to have greater activation elicited by tastants in numerous brain regions commonly reported in gustatory studies, such as the primary somatosensory

area, posterior insula, amygdala, and hypothalamus (Hoogeveen et al., 2015). Relative to their healthy counterparts (mean BMI of 21.4), obese participants (mean BMI of 34.1) had greater activation in the cingulate cortex, insula, orbitofrontal cortex, amygdala, and the striatum (Szalay et al., 2012). If a specific population group is targeted, these factors can be used to filter out unsuitable original studies. If the meta-analyses intended to generalize the results to a broader population, these factors should also be reported, so that readers can understand the composition of the pooled participants and determine if different groups are equally represented. Interested readers may also refer to a recent guideline for conducting food-related neuroimaging experiments, which discussed the common confounders (Smeets et al., 2019).

5. Conclusion

A meta-evaluation has been conducted to evaluate the quality of evidence presented by neuroimaging meta-analyses concerning gustation, food and taste. Results have demonstrated that there exists many meta-analyses on these topics, and they provide moderate to high quality of evidence. There are aspects upon which future meta-analyses in general can be improved. They include accounting for sample overlap, details regarding double checking of data, pre-registration of study protocol, and provision of diagnostics. The current study has compiled a comprehensive list of neuroimaging meta-analyses for gustatory researchers' quick reference. Overall, there seems to be a trend towards a more stringent statistical threshold of cluster-level FWE correction, more meta-analyses conducted and reported in a single paper, and more diversified brain regions reported in MNI space, though the number of studies contained in the meta-analyses remain small to modest. The implementation of GingerALE software in future meta-analyses should ensure that the latest version and cluster-level inference with FWE correction will be used. Using the topic of gustation as an example

of the methodology practice and trends, the current results suggested that similar work is probably needed in other domains of the neuroimaging field. We hope we have demonstrated that this meta-evaluation on taste meta-analyses can be an example for future investigations of trends in literature, specifically in coordinate-based meta-analyses in neuroimaging studies.

Funding: This work received no funding.

Conflict of interest: The authors declared no conflict of interest.

References

Barrós-Loscertales, A., González, J., Pulvermüller, F., Ventura-Campos, N., Bustamante, J.C., Costumero, V., Parcet, M.A., Ávila, C., 2011. Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. Cerebral Cortex 22, 2554-2563.

Bijle, M.N.A., Yiu, C.K.Y., Ekambaram, M., 2018. Calcium-based caries preventive agents: A meta-evaluation of systematic reviews and meta-analysis. Journal of Evidence Based Dental Practice 18, 203-217.

Brett, M., Christoff, K., Cusack, R., Lancaster, J., 2001. Using the Talairach atlas with the MNI template. Neuroimage 13, 85-85.

Brooks, S.J., Cedernaes, J., Schiöth, H.B., 2013. Increased prefrontal and parahippocampal activation with reduced dorsolateral prefrontal and insular cortex activation to food images in obesity: a meta-analysis of fMRI studies. PloS One 8, e60393.

Brooks, S.J., Owen, G., Uher, R., Friederich, H.-C., Giampietro, V., Brammer, M., Williams, S.C., Schiöth, H.B., Treasure, J., Campbell, I.C., 2011. Differential neural responses to food images in women with bulimia versus anorexia nervosa. PloS One 6, e22259.

Brown, S., Gao, X., Tisdelle, L., Eickhoff, S.B., Liotti, M., 2011. Naturalizing aesthetics: brain areas for aesthetic appraisal across sensory modalities. Neuroimage 58, 250-258.

Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience 14, 365-376.

Carp, J., 2012a. On the plurality of (methodological) worlds: estimating the analytic flexibility of fMRI experiments. Frontiers in Neuroscience 6, 149.

Carp, J., 2012b. The secret lives of experiments: methods reporting in the fMRI literature. Neuroimage 63, 289-300.

Chambers, C.D., 2013. Registered reports: a new publishing initiative at Cortex. Cortex 49, 609-610.

Chambers, C.D., Forstmann, B., Pruszynski, J.A., 2017. Registered reports at the European journal of neuroscience: Consolidating and extending peer‐reviewed study pre‐registration. European Journal of Neuroscience 45, 627-628.

Chumbley, J.R., Friston, K.J., 2009. False discovery rate revisited: FDR and topological inference using Gaussian random fields. Neuroimage 44, 62-70.

Cieslik, E.C., Mueller, V.I., Eickhoff, C.R., Langner, R., Eickhoff, S.B., 2015. Three key regions for supervisory attentional control: evidence from neuroimaging meta-analyses. Neuroscience and Biobehavioral Reviews 48, 22-34.

Devoto, F., Zapparoli, L., Bonandrini, R., Berlingeri, M., Ferrulli, A., Luzi, L., Banfi, G., Paulesu, E., 2018. Hungry brains: A meta-analytical review of brain activation imaging

studies on food perception and appetite in obese individuals. Neuroscience and Biobehavioral Reviews 94, 271-285.

Egger, M., Smith, G.D., Sterne, J.A., 2001. Uses and abuses of meta-analysis. Clinical Medicine 1, 478-484.

Eickhoff, S.B., Laird, A.R., Fox, P.M., Lancaster, J.L., Fox, P.T., 2017. Implementation errors in the GingerALE Software: description and recommendations. Human Brain Mapping 38, 7-11.

Eickhoff, S.B., Laird, A.R., Grefkes, C., Wang, L.E., Zilles, K., Fox, P.T., 2009. Coordinate‐based activation likelihood estimation meta‐analysis of neuroimaging data: A random‐effects approach based on empirical estimates of spatial uncertainty. Human Brain Mapping 30, 2907-2926.

Eickhoff, S.B., Nichols, T.E., Laird, A.R., Hoffstaedter, F., Amunts, K., Fox, P.T., Bzdok, D., Eickhoff, C.R., 2016. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 137, 70-85.

Evans, A.C., Collins, D.L., Mills, S., Brown, E., Kelly, R., Peters, T.M., 1993. 3D statistical neuroanatomical models from 305 MRI volumes. Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record. IEEE, pp. 1813-1817.

García‐García, I., Horstmann, A., Jurado, M.A., Garolera, M., Chaudhry, S.J., Margulies, D.S., Villringer, A., Neumann, J., 2014. Reward processing in obesity, substance addiction and non‐substance addiction. Obesity Reviews 15, 853-869.

Garrison, J.R., Done, J., Simons, J.S., 2017. Interpretation of published meta-analytical studies affected by implementation errors in the GingerALE software. Neuroscience and Biobehavioral Reviews [Epub ahead of print], doi: 10.1016/j.neubiorev.2017.1010.1027.

Gentili, C., Messerotti Benvenuti, S., Lettieri, G., Costa, C., Cecchetti, L., 2019. ROI and phobias: The effect of ROI approach on an ALE meta‐analysis of specific phobias. Human Brain Mapping 40, 1814-1828.

Haase, L., Cerf-Ducastel, B., Murphy, C., 2009. Cortical activation in response to pure taste stimuli during the physiological states of hunger and satiety. Neuroimage 44, 1008-1021.

Han, J.E., Boachie, N., Garcia-Garcia, I., Michaud, A., Dagher, A., 2018. Neural correlates of dietary self-control in healthy adults: A meta-analysis of functional brain imaging studies. Physiology and Behavior 192, 98-108.

Hoogeveen, H.R., Dalenberg, J.R., Renken, R.J., ter Horst, G.J., Lorist, M.M., 2015. Neural processing of basic tastes in healthy young and older adults—an fMRI study. Neuroimage 119, 1-12.

Huerta, C.I., Sarkar, P.R., Duong, T.Q., Laird, A.R., Fox, P.T., 2014. Neural bases of food perception: Coordinate‐based meta‐analyses of neuroimaging studies in multiple modalities. Obesity 22, 1439-1446.

Kennedy, J., Dimitropoulos, A., 2014. Influence of feeding state on neurofunctional differences between individuals who are obese and normal weight: a meta-analysis of neuroimaging studies. Appetite 75, 103-109.

Kurth, F., Zilles, K., Fox, P.T., Laird, A.R., Eickhoff, S.B., 2010. A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. Brain Structure and Function 214, 519-534.

Lancaster, J.L., Tordesillas‐Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J.C., Fox, P.T., 2007. Bias between MNI and Talairach coordinates analyzed using the ICBM‐152 brain template. Human Brain Mapping 28, 1194-1205.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Medicine 6, e1000100.

Müller, V.I., Cieslik, E.C., Laird, A.R., Fox, P.T., Radua, J., Mataix-Cols, D., Tench, C.R., Yarkoni, T., Nichols, T.E., Turkeltaub, P.E., 2018. Ten simple rules for neuroimaging meta-analysis. Neuroscience and Biobehavioral Reviews 84, 151-161.

McNorgan, C., 2012. A meta-analytic review of multisensory imagery identifies the neural correlates of modality-specific and modality-general imagery. Frontiers in Human Neuroscience 6, 285.

Moher, D., Cook, D.J., Eastwood, S., Olkin, I., Rennie, D., Stroup, D.F., 2000. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Oncology Research and Treatment 23, 597-602.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of Internal Medicine 151, 264-269.

Nakagawa, S., Noble, D.W., Senior, A.M., Lagisz, M., 2017. Meta-evaluation of meta-analysis: ten appraisal questions for biologists. BMC Biology 15, 18.

Ng, J., Stice, E., Yokum, S., Bohon, C., 2011. An fMRI study of obesity, food reward, and perceived caloric density. Does a low-fat label make food less appealing? Appetite 57, 65-72.

Noori, H.R., Linan, A.C., Spanagel, R., 2016. Largely overlapping neuronal substrates of reactivity to drug, gambling, food and sexual cues: a comprehensive meta-analysis. European Neuropsychopharmacology 26, 1419-1430.

Panic, N., Leoncini, E., De Belvis, G., Ricciardi, W., Boccia, S., 2013. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. PloS One 8, e83138.

Peters, J.P., Hooft, L., Grolman, W., Stegeman, I., 2015. Reporting quality of systematic reviews and meta-analyses of otorhinolaryngologic articles based on the PRISMA Statement. PloS One 10, e0136540.

Poldrack, R.A., Baker, C.I., Durnez, J., Gorgolewski, K.J., Matthews, P.M., Munafò, M.R., Nichols, T.E., Poline, J.-B., Vul, E., Yarkoni, T., 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. Nature Reviews Neuroscience 18, 115-126.

Pollock, M., Fernandes, R.M., Hartling, L., 2017. Evaluation of AMSTAR to assess the methodological quality of systematic reviews in overviews of reviews of healthcare interventions. BMC Medical Research Methodology 17, 48.

Pursey, K.M., Stanwell, P., Callister, R.J., Brain, K., Collins, C.E., Burrows, T.L., 2014. Neural responses to visual food cues according to weight status: a systematic review of functional magnetic resonance imaging studies. Frontiers in nutrition 1, 7.

Radua, J., Mataix-Cols, D., Phillips, M.L., El-Hage, W., Kronhaus, D., Cardoner, N., Surguladze, S., 2012. A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps. European Psychiatry 27, 605-611.

Reddan, M.C., Lindquist, M.A., Wager, T.D., 2017. Effect size estimation in neuroimaging. JAMA psychiatry 74, 207-208.

Rothemund, Y., Preuschhof, C., Bohner, G., Bauknecht, H.-C., Klingebiel, R., Flor, H., Klapp, B.F., 2007. Differential activation of the dorsal striatum by high-calorie visual food stimuli in obese individuals. Neuroimage 37, 410-421.

Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D., Nichols, T.E., 2009. Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. Neuroimage 45, 810-823.

Santel, S., Baving, L., Krauel, K., Münte, T.F., Rotte, M., 2006. Hunger and satiety in anorexia nervosa: fMRI during cognitive processing of food pictures. Brain Research 1114, 138-148.

Satpute, A.B., Kang, J., Bickart, K.C., Yardley, H., Wager, T.D., Barrett, L.F., 2015. Involvement of sensory regions in affective experience: a meta-analysis. Frontiers in Psychology 6, 1860.

Sescousse, G., Caldú, X., Segura, B., Dreher, J.-C., 2013. Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. Neuroscience and Biobehavioral Reviews 37, 681-696.

Simmons, J.P., Nelson, L.D., Simonsohn, U., 2011. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological Science 22, 1359-1366.

Small, D.M., Gregory, M.D., Mak, Y.E., Gitelman, D., Mesulam, M.M., Parrish, T., 2003. Dissociation of neural representation of intensity and affective valuation in human gustation. Neuron 39, 701-711.

Smeets, P.A., Dagher, A., Hare, T.A., Kullmann, S., van der Laan, L.N., Poldrack, R.A., Preissl, H., Small, D., Stice, E., Veldhuizen, M.G., 2019. Good practice in food-related neuroimaging. American Journal of Clinical Nutrition 109, 491-503.

Stevinson, C., Lawlor, D., 2004. Searching multiple databases for systematic reviews: added value or diminishing returns? Complementary Therapies in Medicine 12, 228-232.

Stice, E., Spoor, S., Bohon, C., Small, D., 2008. Relation between obesity and blunted striatal response to food is moderated by TaqIA A1 allele. Science 322, 449-452.

Szalay, C., Aradi, M., Schwarcz, A., Orsi, G., Perlaki, G., Németh, L., Hanna, S., Takács, G., Szabó, I., Bajnok, L., 2012. Gustatory perception alterations in obesity: an fMRI study. Brain Research 1473, 131-140.

Tahmasian, M., Zarei, M., Noori, K., Khazaie, H., Samea, F., Spiegelhalder, K., Eickhoff, S.B., Van Someren, E., Eickhoff, C.R., 2018. Reply to Hua Liu, HaiCun Shi and PingLei Pan: Coordinate based meta-analyses in a medium sized literature: Considerations, limitations and road ahead. Sleep Medicine Reviews 42, 236.

Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. Thieme, New York.

Tang, D., Fellows, L., Small, D., Dagher, A., 2012. Food and drug cues activate similar brain regions: a meta-analysis of functional MRI studies. Physiology and Behavior 106, 317-324.

Turkeltaub, P.E., Eickhoff, S.B., Laird, A.R., Fox, M., Wiener, M., Fox, P., 2012. Minimizing within‐experiment and within‐group effects in activation likelihood estimation meta‐analyses. Human Brain Mapping 33, 1-13.

van der Laan, L.N., De Ridder, D.T., Viergever, M.A., Smeets, P.A., 2011. The first taste is always with the eyes: a meta-analysis on the neural correlates of processing visual food cues. Neuroimage 55, 296-303.

van der Laan, L.N., Smeets, P.A., 2015. You are what you eat: a neuroscience perspective on consumers' personality characteristics as determinants of eating behavior. Current Opinion in Food Science 3, 11-18.

van Meer, F., van der Laan, L.N., Adan, R.A., Viergever, M.A., Smeets, P.A., 2015. What you see is what you eat: an ALE meta-analysis of the neural correlates of food viewing in children and adolescents. Neuroimage 104, 35-43.

Veldhuizen, M.G., Albrecht, J., Zelano, C., Boesveldt, S., Breslin, P., Lundström, J.N., 2011. Identification of human gustatory cortex by activation likelihood estimation. Human Brain Mapping 32, 2256-2266.

Verhagen, J.V., Engelen, L., 2006. The neurocognitive bases of human multimodal food perception: sensory integration. Neuroscience and Biobehavioral Reviews 30, 613-650.

Yeung, A.W.K., 2017. Do neuroscience journals accept replications? A survey of literature. Frontiers in Human Neuroscience 11, 468.

Yeung, A.W.K., 2018. Sex differences in brain responses to food stimuli: a meta‐analysis on neuroimaging studies. Obesity Reviews 19, 1110-1115.

Yeung, A.W.K., Goto, T.K., Leung, W.K., 2017. Basic taste processing recruits bilateral anteroventral and middle dorsal insulae: an activation likelihood estimation meta-analysis of fMRI studies. Brain and Behavior 7, e00655.

Yeung, A.W.K., Goto, T.K., Leung, W.K., 2018. Affective value, intensity and quality of liquid tastants/food discernment in the human brain: An activation likelihood estimation meta-analysis. Neuroimage 169, 189-199.

Zhu, Y., Hu, X., Wang, J., Chen, J., Guo, Q., Li, C., Enck, P., 2012. Processing of food, body and emotional stimuli in anorexia nervosa: a systematic review and meta‐analysis of functional magnetic resonance imaging studies. European Eating Disorders Review 20, 439-450.

Figure legends

Figure 1. Modified PRISMA flow diagram showing the procedure of the meta-analysis selection process for meta-evaluation. For Scopus and Web of Science, the title, abstract and keywords of the publication records were searched; for PsycINFO and PubMed, no search restrictions were placed. All identified meta-analyses reporting brain activations related to gustatory, food or taste stimuli were screened.

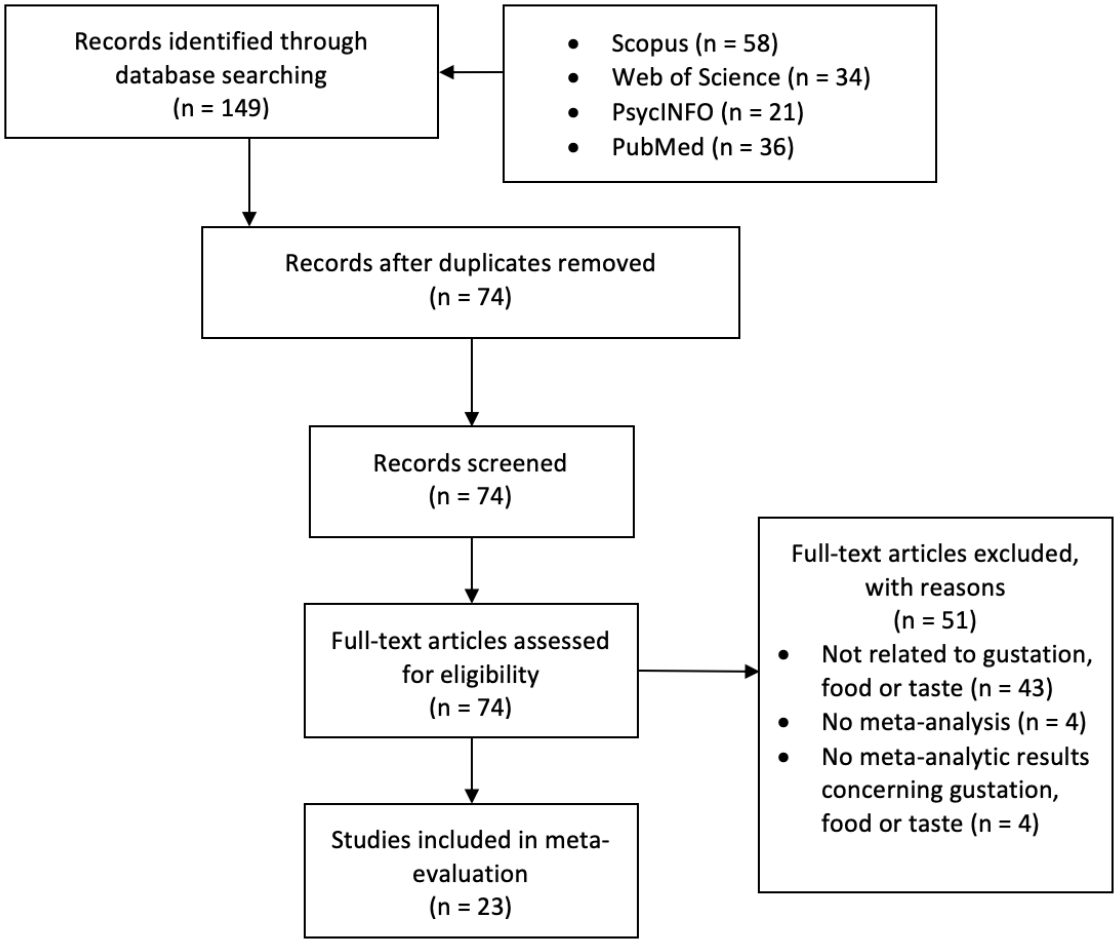

Figure 2. Time trend of brain space and statistical threshold selected. The numbers represent the number of meta-analytic studies. (A) Meta-analyses are increasingly reporting the locations of brain activations in standard Montreal Neurological Institute (MNI) coordinate space, and (B) adopting the cluster-level inference using familywise error rate (FWE) correction. FDR, false discovery rate.

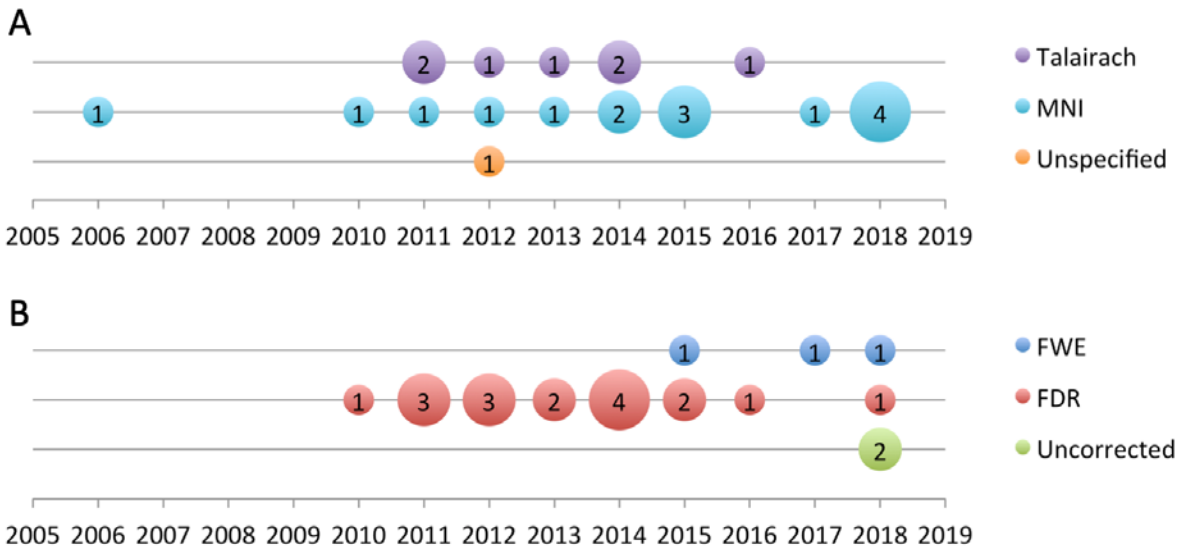

Figure 3. (A) Time trend of number of original papers included in each meta-analysis. (B) Time trend of number of relevant meta-analyses conducted in each paper. (C) Time trend of number of number of foci reported in each paper.

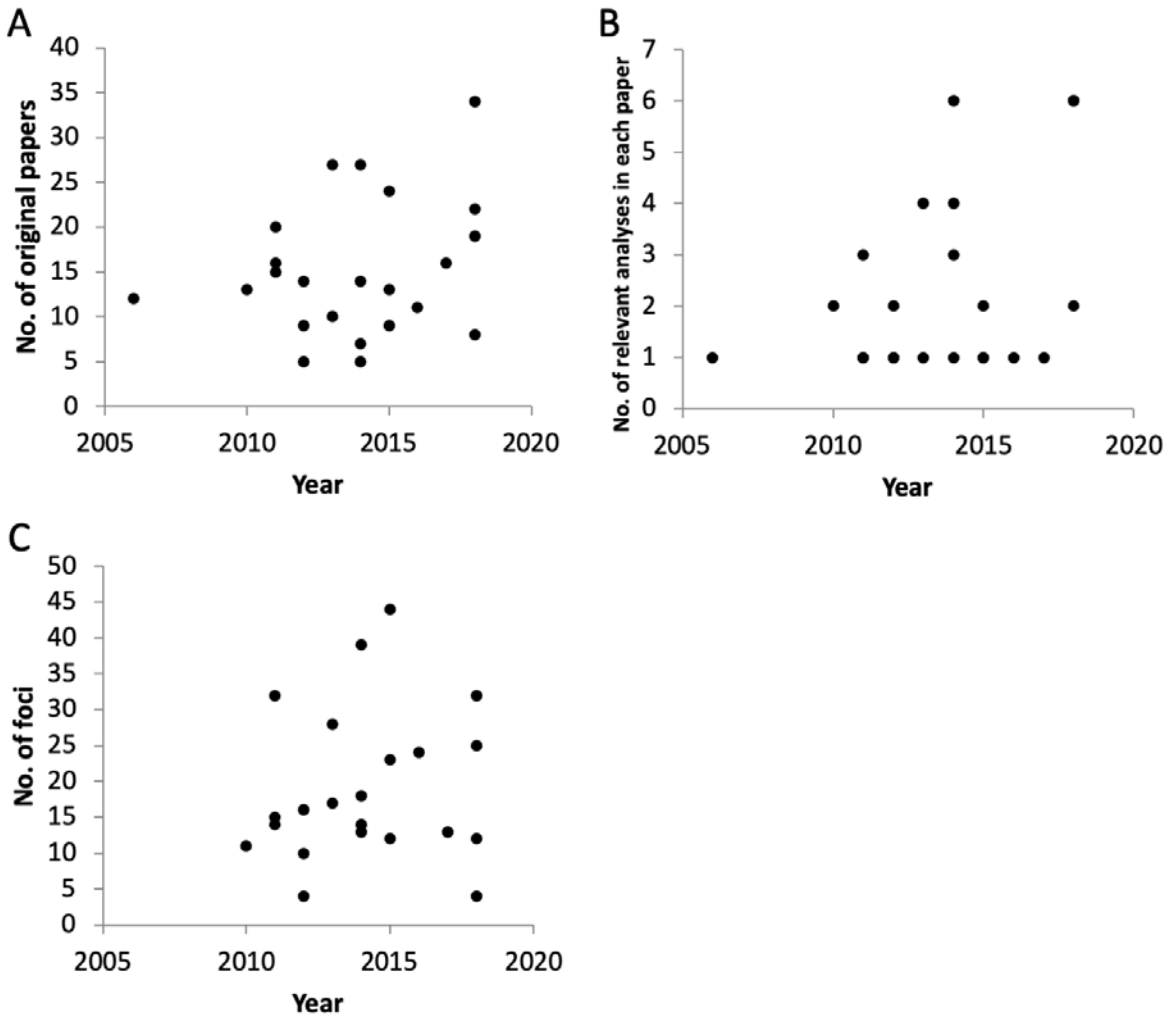

Figure 4. The prevalence of different statistical thresholds adopted by the surveyed meta-analytic studies.

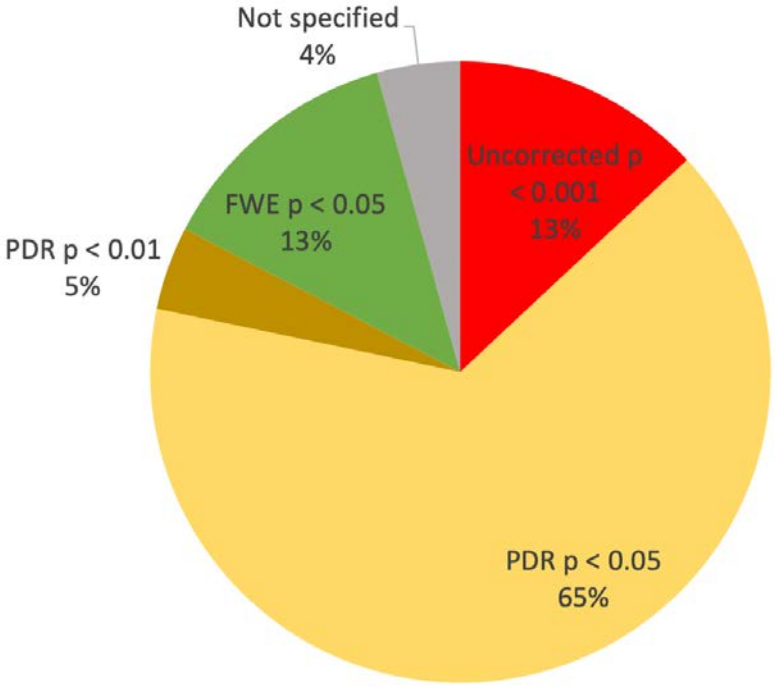

Table 1. Descriptive data of the 23 meta-analyses analyzed.

	Mean age	Mean BMI	Stimuli	Outcome	Number of relevant meta-analyses performed	Number of null results	Number of original papers analyzed	Number of foci reported	Statistical threshold[b]	Where the results were presented	Results in brain space	Research relevance
	1=<18 y; 2=>18 y; 3=NA[a]	1=<18.5; 2=18.5-24.9; 3=25-29.9; 4=>30; 5=NA[a]	1=visual; 2=tastant; 3=real food; 4=mental imagery; 5=others	1=within-group; 2=between-group					1=FWE; 2=FDR; 3=uncorrected (unc)		1=Talairach; 2=MNI; 3=NA[a]	1=coordinates (listed in a table); 2=downloadable brain map (in NIFTI format)
Devoto et al. (2018)	2	2, 4	1, 2, 3, 4	1, 2	6	0	22	12	2, 3	Table 3	2	1, 2
Han et al. (2018)	1, 2	2, 3, 4	1	1, 2	6	0	19	32	3	Tables 3-5	2	1, 2
Yeung (2018)	1, 2	2, 3, 4	1, 2, 3	2	2	0	8	4	3 (+ cluster > 200mm^3)	Table 2	2	1
Yeung et al. (2018)	1, 2	1, 2, 3, 4	2, 3	1	6	1	34	25	1	Tables 2-3	2	1, 2
Yeung et al. (2017)	2	2, 3	2	1	1	0	16	13	1	Table 2	2	1
Noori et al. (2016)	2	1, 2, 3, 4	1, 4	1	1	0	11	24	2 (+ cluster > 200mm^3)	eTable 9	1	1
Satpute et al. (2015)	3	5	2, 3	1	1	0	9	23	1	Data sheet 2	2	1
van der Laan and Smeets (2015)	1, 2	2, 4	1, 3	1, 2	1	0	13	12	2 (+ cluster > 100mm^3)	Table 1	2	1
van Meer et al. (2015)	1, 2	2	1, 4	1	2	0	24	44	2 (+ cluster > 100mm^3)	Tables 2-3	2	1
García-García et al. (2014)	1, 2	2, 3, 4	1, 3	2	4	0	14	14	2 (+ cluster > 200mm^3)	Tables 2, S7	2	1
Huerta et al. (2014)	3	5	1, 2, 3	1	3	0	27	39	2 (+ cluster > 200mm^3)	Tables 4-6	1	1
Kennedy and Dimitropoulos (2014)	3	2, 3, 4	1	2	6	1	7	18	2 (+ cluster > 100mm^3)	Tables 2-3	1	1
Pursey et al. (2014)	3	5	1	1	1	0	5	13	2 (+ cluster > 100mm^3)	Table S4	2	1
Brooks et al. (2013)	1, 2	5	1, 2, 3	2	4	0	10	17	2 (+ cluster > 100mm^3)	Tables 2-3	1	1
Sescousse et al. (2013)	2	5	2, 3	1	1	0	27	28	2 (FDR $p < 0.01$ + cluster > 600mm^3)	Table S3	2	1
McNorgan (2012)	3	5	4	1	1	0	5	10	2	Table 3	2	1
Tang et al. (2012)	2	2	1	1	1	0	14	16	2	Table 4	3	1
Zhu et al. (2012)	3	1, 2	1, 2, 3	2	2	0	9	4	2 (+ cluster > 200mm^3)	Table 2	1	1

Brown et al. (2011)	3	5	2, 3	1	1	0	16	14	2 (+ cluster > 10 voxels)	Table S5	1	1
van der Laan et al. (2011)	3	2	1	1	3	0	20	32	2 (+ cluster > 100mm^3)	Tables 2-3	2	1
Veldhuizen et al. (2011)	2	5	2	1	1	0	15	15	2 (+ cluster > 100mm^3)	Table 2	1	1
Kurth et al. (2010)	3	5	2	1	2	0	13	11	2	Table 3	2	1
Verhagen and Engelen (2006)	2	5	2	1	1	0	12	.	.	Figure 3	2	.

[a] NA, not reported.

[b] FWE, familywise error rate. FDR, false discovery rate. Unless otherwise specified, FWE refers to cluster-level FWE $p < 0.05$, with cluster forming threshold of uncorrected $p < 0.001$. Meanwhile, FDR and uncorrected refer to voxel-level thresholds, and FDR refers to FDR $p < 0.05$ and uncorrected refers to uncorrected $p < 0.001$, unless otherwise specified.

Table 2. Quality evaluation of the included meta-analyses.

	Defined research question (aspects)	Systematic literature search (databases)	Defined inclusion and exclusion criteria	Accounting for sample overlap (method)	Inclusion of studies with whole brain coverage only	Studies converted to a common reference brain space	Double checking of data	Descriptions of included studies in a table	Pre-registration of study protocol	Default methods and parameters of established software	Diagnostics	Total score (1 for each Yes)
		1 = Web of Science 2 = Scopus 3 = PubMed 4 = Google Scholar 5 = Ovid 6 = Science Direct 7 = Cochrane 8 = Neurosynth 9 = BrainMap		1 = Adjusted for within-group effects 2 = Used data from one experiment per subject group		1 = Brett transformation (Brett et al., 2001) 2 = Lancaster transformation (Lancaster et al., 2007)	1 = Data extracted by two or more authors 2 = One author, double check			1 = GingerALE (version, v) 2 = AES-SDM 3 = MKDA	1 = Experiment contributions to significant clusters 2 = Funnel plots 3 = Heterogeneity analysis	
Devoto et al. (2018)	Yes (obese vs healthy; visual vs gustatory; hungry vs satied)	Yes (3)	Yes	Unclear	Yes	Yes (2)	Unclear	Yes (its Table S1)	Unclear	Unclear (1, v2.3.6; CluB[a])	Yes (1)	7
Han et al. (2018)	Yes (food craving regulation vs baseline; its modulation by BMI; food craving regulation vs food decision making)	Yes (3, 4, 5, 6, 8)	Yes	Unclear	Yes	Yes (included studies reporting in MNI[b] space only)	Yes (1)	Yes (its Table 2)	Unclear	Yes (2)	Yes (3)	9
Yeung (2018)	Yes (males vs females)	Yes (2, 3)	Yes	Unclear	Yes	Yes (2)	Unclear	Yes (its Table 1)	Unclear	Yes (1, v2.3.6)	Unclear	7
Yeung et al. (2018)	Yes (affective value; intensity; quality)	Yes (3, 8, 9)	Yes	Unclear	Yes	Yes (2)	Unclear	Yes (its Table 1)	Unclear	Yes (1, v2.3.6)	Yes (1)	8
Yeung et al. (2017)	Yes (taste vs tasteless)	Yes (3, 5)	Yes	Unclear	Yes	Yes (2)	Unclear	Yes (its Table 1)	Unclear	Yes (1, v2.3.6)	Yes (1)	8

Noori et al. (2016)	Yes (food vs neutral)	Yes (3)	Yes	Unclear	No	Yes (2)	Yes (2)	Yes (its eTable 2)	Unclear	Yes (1, v2.0.4)	Unclear	7
Satpute et al. (2015)	Yes (affective food / liquid vs neutral)	Unclear (self manually coded database)	Yes	Unclear	No	Yes (1)	Unclear	No (its Data sheet S1 listed the studies only)	Unclear	Yes (3)	Unclear	4
van der Laan and Smeets (2015)	Yes (brain activation to food stimuli that correlated to personality characteristic scores; restrained eating vs control)	Yes (3)	Yes	Unclear	No	Yes (2?)	Unclear	Yes (its Table S2)	Unclear	Yes (1, v?)	Yes (1)	7
van Meer et al. (2015)	Yes (food vs non-food)	Yes (3, 4)	Yes	Yes (1)	Yes	Yes (2?)	Unclear	Yes (its Table 1)	Unclear	Yes (1, v?)	Yes (1)	9
García-García et al. (2014)	Yes (obese vs healthy)	Yes (3)	Yes	Unclear	Yes	Unclear	Unclear	Yes (its Table S1)	Unclear	Yes (1, v2.3)	Unclear	6
Huerta et al. (2014)	Yes (food vs non-food)	Yes (3, 9)	Yes	Unclear	Yes	Yes (2)	Unclear	Yes (its Tables 1-3)	Unclear	Yes (1, v2.1)	Unclear	7
Kennedy and Dimitropoulos (2014)	Yes (obese vs healthy)	Yes (3, 4, 5, 6)	Yes	Unclear	Yes	Yes (2?)	Unclear	Yes (its Table 1)	Unclear	Yes (1, v?)	Unclear	7
Pursey et al. (2014)	Yes (pre-weight loss vs post-weight loss)	Yes (1, 2, 3, 5, 7, EMBASE, CINAHL, Informit Health Collection, Proquest)	Yes	Unclear	Yes	Yes (2?)	Yes (1)	Yes (its Tables S1-S3)	Unclear	Yes (1, v?)	Unclear	8
Brooks et al. (2013)	Yes (obese / overweight vs healthy)	Yes (1, 3, 4, 5, 6)	Yes	Unclear	Yes	Yes (2?)	Unclear	Yes (its Table 1)	Unclear	Yes (1, v?)	Unclear	7
Sescousse et al. (2013)	Yes (food vs non-food + correlation with pleasantness score)	Yes (3)	Yes	Yes (2)	Yes	Yes (2)	Yes (1)	Yes (its Table 3)	Unclear	Yes (1, v2.2)	Unclear	9

McNorgan (2012)	Yes (gustatory imagery vs baseline)	Yes (3, 4)	Yes	Unclear	Yes	Unclear	Unclear	Yes (its Table 1)	Unclear	Yes (1, v2.1)	Unclear	6
Tang et al. (2012)	Yes (food vs non-food)	Yes (3)	Yes	Unclear	Yes	Unclear	Unclear	Yes (its Table 1)	Unclear	Yes (1, v2.1.1)	Unclear	6
Zhu et al. (2012)	Yes (anorexic vs healthy)	Yes (1, 3, 4, 5, 7)	Yes	Unclear	Yes (ROI[c] studies were initially included, but results were unchanged if they were excluded)	Yes (2)	Yes (1)	Yes (its Table 1)	Unclear	Yes (1, v2.0)	Unclear	8
Brown et al. (2011)	Yes (affective value)	Yes (1)	Yes	Unclear	Yes	Yes (2?)	Yes (1)	Yes (its Table S2)	Unclear	Yes (1, v2.0.1a3)	Unclear	8
van der Laan et al. (2011)	Yes (food vs non-food; hungry vs satied; high vs low energy food)	Yes (3)	Yes	Unclear	Yes	Yes (2?)	Unclear	Yes (its Table 1)	Unclear	Yes (1, v?)	Yes (1)	8
Veldhuizen et al. (2011)	Yes (taste vs tasteless)	Yes (3, 5)	Yes	Unclear	Unclear	Yes (2?)	Unclear	Yes (its Table 1)	Unclear	Yes (1, v2.0.1)	Unclear	6
Kurth et al. (2010)	Yes (taste vs tasteless)	Yes (3, 9)	Yes	Unclear	Yes	Unclear	Unclear	Yes (its Supplement 1)	Unclear	Yes (1, v?)	Unclear	6
Verhagen and Engelen (2006)	Yes (taste vs tasteless)	Unclear	Yes	Yes (2)	Unclear	Yes (1)	Unclear	Yes (its Table 1)	Unclear	Unclear (only plotted all coordinates extracted from included studies onto a standard brain template “ch2” in MNI space	Unclear	5

[a] CluB (Clustering the Brain) software is available at https://goo.gl/rB2DQx and is currently in beta version.
[b] MNI, Montreal Neurological Institute.
[c] ROI, region-of-interest studies.