

# Neurometabolite differences in Autism as assessed with Magnetic Resonance Spectroscopy: a systematic review and meta-analysis.

Alice R. Thomson<sup>1,2</sup>, Duanghathai Pasanta<sup>1</sup>, Tomoki Arichi<sup>2,3</sup>, and Nicolaas A. Puts<sup>1,2</sup>

<sup>1</sup>Department of Forensic and Neurodevelopmental Sciences, King's College London

<sup>2</sup>MRC Centre for Neurodevelopmental Disorders, King's College London

<sup>3</sup>Centre for the Developing Brain, King's College London, London,

## Corresponding author:

Nicolaas A. Puts

Department of Forensic and Neurodevelopmental Sciences  
Institute of Psychiatry, Psychology & Neuroscience (IoPPN)  
King's College London  
E1.05, 16 De Crespigny Park, London, SE5 8AF

[nicolaas.puts@kcl.ac.uk](mailto:nicolaas.puts@kcl.ac.uk)

## Abstract

<sup>1</sup>H-Magnetic Resonance Spectroscopy (MRS) is a non-invasive technique that can be used to quantify the concentrations of metabolites in the brain in vivo. MRS findings in the context of autism are inconsistent and conflicting. We performed a *systematic review and meta-analysis* of MRS studies observing glutamate and gamma-aminobutyric acid (GABA), as well as brain metabolites involved in energy metabolism (glutamine, creatine), neural integrity (e.g. n-acetyl aspartate (NAA), choline, myo-inositol) and neuro-inflammation (glutathione) in autism cohorts. Data were extracted and grouped by metabolite, brain region and several other factors before calculation of standardised effect sizes. Overall, we find significantly lower concentrations of GABA and NAA in autism, indicative of disruptions to the balance between excitation/inhibition within brain circuits, as well as neural integrity. Further analysis found these alterations are most pronounced in autistic children and in limbic brain regions relevant to autism phenotypes. Additionally, we show how study outcome varies due to demographic and methodological factors (e.g. medication use), emphasising the importance of conforming with consensus study designs and transparent reporting.

## Key words

Autism, edited MRS, GABA, Glutamate, E/I balance, metabolites, neurodevelopment.

## Highlights

- We performed a meta-analysis of *Magnetic Resonance Spectroscopy* findings in autism.
- We find significantly lower concentrations of GABA and NAA in children with autism.
- These alterations were most pronounced in limbic brain regions.
- Demographic factors (including sex) contribute to MRS study outcome variation.

## 1. Introduction

### 1.1 background

Autism Spectrum Condition (henceforth referred to as ‘autism’ as preferred by the autistic community) is a neurodevelopmental condition, affecting 1-2% of the population and characterised by differences in cognition, social communication, and repetitive and restricted interests (American Psychiatric Association, 2000, 2013; Horder et al., 2018). Sensory differences are also extremely common and are reported in 60-95% of autistic individuals (He et al., 2021; Klintwall et al., 2011; Rogers & Ozonoff, 2005).

Most cases of autism are idiopathic in nature (Horder et al., 2018) with around a thousand gene variants linked to the condition in humans (Grove et al., 2019). Stratification of patient aetiologies based on shared neurobiological ‘endpoints’ of this genetic heterogeneity is essential to advance our understanding of the neurodevelopmental abnormalities that underpin autism at the biochemical, neuronal and circuit level (Brian et al., 2016; Sanders, 2015; Wolfers et al., 2019)

One such common ‘endpoint’ across autism aetiologies might be a disruption in the balance of excitation/inhibition in brain circuits (E/I; Rubenstein & Merzenich, 2003). Typical brain states rely on a homeostasis between E/I, predominantly determined by the endogenous activity of glutamate (Glu) and gamma-aminobutyric acid (GABA), the adult human brain’s principal excitatory and inhibitory neurotransmitters respectively. The maintenance of this E/I balance is a critical component of typical neurodevelopment and neuronal computation (Ghisleni et al., 2015; Rubenstein & Merzenich, 2003; S. Zhou & Yu, 2018).

In autism, an 'E/I' imbalance within brain circuits is thought to lead to differences in sensory function, memory, and social and non-social cognition (Rubenstein & Merzenich, 2003).

Several lines of evidence support this theory, including genetic, preclinical animal, and post-mortem studies of human brain tissue (Ajram et al., 2019; DeLorey et al., 2008; Etherton et al., 2011; Fatemi et al., 2002; Hashemi et al., 2017; Moreno-De-Luca et al., 2013; Port et al., 2019; Tuchman & Rapin, 2002; Yan et al., 2005). For example, studies have found various autism-associated genes with a role in neuronal excitatory and inhibitory cascades (Grove et al., 2019; Havdahl et al., 2021; Moreno-De-Luca et al., 2013). Other studies have found differences in inhibitory interneuron availability and synaptic pruning (Hashemi et al., 2017; Kim et al., 2017; Tang et al., 2014), and many animal models of autism show differences in E/I signalling while showing autistic phenotypes (DeLorey et al., 2008; Etherton et al., 2011; Qi et al., 2022). Finally, seizure disorders, thought to arise due to abnormally increased excitation brain circuits (Rubenstein & Merzenich, 2003) regularly co-occur with autism (ranging from 5-40%; Tuchman & Rapin, 2002).

However, other biological theories exist; the redox hypothesis is an additional (although not necessarily discrete) aetiological model of autism in which environmentally induced oxidative stress leads to altered neurodevelopment in genetically vulnerable individuals (Endres et al., 2017). Consistent with this, post-mortem tissue samples from autistic individuals show evidence of an abnormal redox state and glutathione (GSH) deficiency (Adams et al., 2009; Chauhan et al., 2012; F. Gu et al., 2013). GSH is a cellular antioxidant which donates electrons for the reduction of reactive oxygen species (ROS e.g hydrogen peroxide) within neurons and glial cells (Dringen, 2000) and thus acts as a crucial defence against neuronal oxidative stress (Dringen, 2000). Preclinical researchers also find reductions in GSH brain concentrations in

mouse models of autism, with further GSH depletion increasing the occurrence of autism-like behaviours (Nadeem et al., 2019). (Bjørklund et al., 2020; Frye et al., 2013).

*<sup>1</sup>H-Magnetic Resonance Spectroscopy (MRS)* is a non-invasive technique that can be used to quantify the concentrations of endogenous brain metabolites including GSH, Glu and GABA in the human brain in vivo (Puts & Edden, 2012). As proxy measures of E/I balance (Glu/GABA) and neuroinflammation (GSH), MRS is therefore highly valuable for validating the E/I imbalance/redox hypothesis of autism in humans and so identifying human biochemical pathways for intervention.

Beyond GABA/Glu and GSH, other MRS measured metabolites have been implicated in the pathophysiology of autism, including those involved in energy metabolism (glutamine (Gln) and creatine (Cr)), and neural cytoarchitecture and integrity (e.g. N-acetyl aspartate (NAA), choline (Cho), myo-inositol (mi)); Levitt et al., 2003; Libero et al., 2016; O'Neill et al., 2020; Oner et al., 2007; Vasconcelos et al., 2008), although preclinical evidence complementing this work is more limited.

## **1.2 MRS**

MRS works by suppressing the brain water signal to detect other <sup>1</sup>H-proton containing metabolites at lower concentrations (in the order of 0 – 20 mM). <sup>1</sup>H protons in different metabolites have different chemical environments (e.g. surrounded by more electrons than others) giving each metabolite a unique resonant (& spatial) frequency measured in parts per million (ppm). This spatial frequency is plotted along a chemical shift axis. The amplitude of

each metabolite peak along this axis is proportional to the number of  $^1\text{H}$  protons detected and as such the metabolite concentration (Puts & Edden, 2012). To obtain an acceptable signal to noise ratio (SNR), metabolites are quantified from an average MRS spectrum, the result of many repeated measurements (transients) from the same voxel.

Conventional MR pulse sequences for MRS acquisition, including Point Resolved Spectroscopy (PRESS; Bottomley, 1987) and Stimulated Echo Acquisition Mode (STEAM; Frahm et al., 1987), can detect and so quantify high-concentration metabolites including NAA, total choline and total creatine. Tailored 'edited' MRS sequences are required to resolve GABA, Glu and GSH, as their signals are usually masked by metabolites in greater concentrations (e.g. NAA) and so with stronger resonances (Mullins et al., 2014). One such edited approach is MESHcher-GARwood Point RESolved Spectroscopy sequence (MEGA-PRESS; Mescher et al., 1996, 1998). MEGA-PRESS uses a 1.9 ppm editing pulse to selectively modulate the GABA signal during the acquisition phase (edit-ON). Subtraction of this edited spectra (edit-ON) from an unedited spectrum (edit-OFF) gives the difference (DIFF) spectrum, which contains metabolite frequencies modulated by the GABA-selective editing, including GABA at 3.02 ppm. A similar approach can be used for GSH by placing editing-ON pulses at 4.56 ppm. Due to contamination with macromolecules which are also affected by the GABA-selective editing pulse, GABA signal is often denoted as GABA+ (GABA + macromolecules; Mullins et al., 2014). Similarly, due to the poor chemical shift dispersion achieved at 1.5-3 Tesla (overlap of metabolite spatial frequencies on the chemical shift axis) it is often not possible to distinguish Glu signals from that of its precursor glutamine (Gln). Glx therefore describes the combined concentrations of Gln and Glu (Mullins et al., 2014).

### **1.3 The current study**

While MRS has previously been used to identify autism-associated alterations in brain metabolite concentrations, variation in brain region, <sup>1</sup>H-MRS parameters, and small sample sizes have limited interpretation, with highly conflicting outcomes.

Here, we quantitatively and systemically review previous MRS work in autism, grouping results by brain region and, as metabolite concentrations have been shown to vary across the lifespan, by age (Ghisleni et al., 2015; Porges et al., 2021; Reyngoudt et al., 2012). To extensively cover the previous literature, we focus on seven key metabolites; GABA, Glu/Glx, Gln, NAA, creatine, choline and ml, these being the most widely studied metabolites in autism (Ford & Crewther, 2016a). We aimed to comprehensively summarise the MRS-based evidence for autism associated alterations in these brain metabolite concentrations and as such, autism aetiological theories. We performed appropriate assessments for MRS data quality, which have not yet been performed in the autism field, and controlled for non-independence of data, as most studies report more than one metabolite and/or measure from more than one brain region per participant. We additionally investigate potential sources of bias within the included studies such as MRS data acquisition parameters, demographic variables, and quality of MRS data. In doing so we aimed to isolate factors that contributed to the heterogeneity of previous findings and as such guide researchers towards more replicable study designs. Finally, we discuss the potential source of metabolite differences in autism at the mechanistic level.

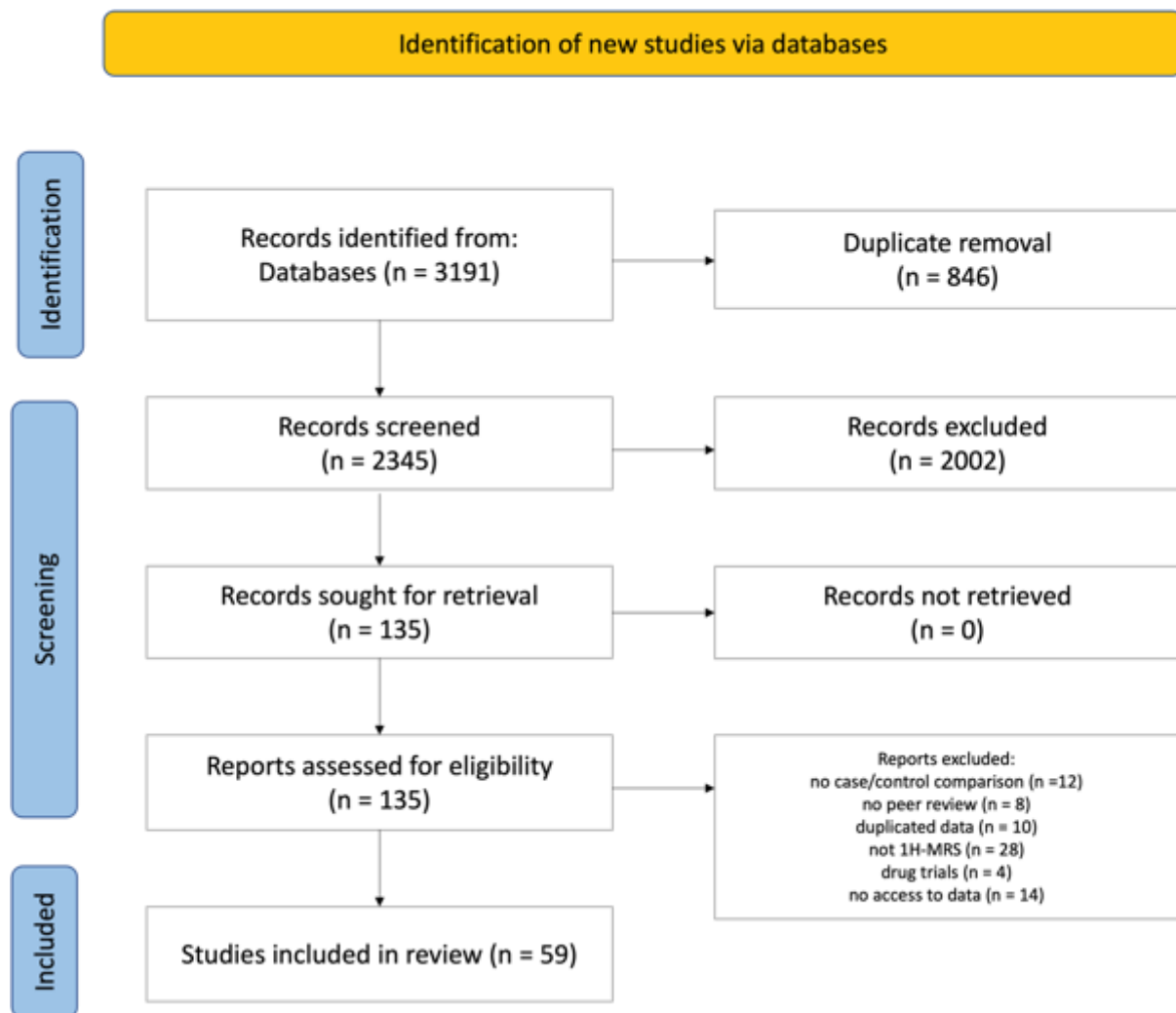
## **2. Materials and methods**

## 2.1 Search criteria and study selection strategy

A systematic search of databases (Ovid Medline, Pubmed, Web of Science and Google Scholar) was performed using the search terms selected using the Boolean method from *litserchr* in R (29/09/22), these being; (“ASD” OR “Autism Spectrum Disorder” OR “Autistic disorder” OR “Asperger's Syndrome” OR “Autism”) AND (“<sup>1</sup>H-MRS” OR “in vivo magnetic resonance spectroscopy” OR “magnetic resonance spectroscopy” OR “proton magnetic resonance spectroscopy” OR “MRS” OR “GABA” OR “Glutamate” OR “Creatine” OR “Choline” OR “metabolite”). Duplicate papers were removed before abstract screening using the *metagear* package in R (Lajeunesse, 2016). **Papers were included if they met all of the following criteria; (1) use of in-vivo MRS to measure metabolites in the brain; (2) investigation of brain metabolite concentrations in autism and non-autism groups (3) MRS measured from a specified brain region and (4) the study was published in a peer-reviewed journal, and was written in English or translated to English.** Relevant articles from the reference sections of included studies were also identified and added to the analysis. Historically, autistic individuals have been diagnosed with autistic disorder (AD) or Asperger’s syndrome (AS) depending on the severity of the autistic phenotype (Colleen & Stone, 2014). Present classification criteria make this subgrouping redundant, placing both groups under autism spectrum disorder (ASD). As several MRS autism studies predate this change in diagnostic approach, we included studies reporting cohorts of ASD, AD and AS diagnosis in the meta-analysis which explores autism. The impact of the explicit diagnosis will however be explored (ASD/AD/AS).



The initial search returned 3191 studies (29/09/22). 2345 studies were eligible for abstract screening after removal of duplicates. 2002 studies were excluded in the abstract screening (with full text screening performed if necessary). Six relevant articles from the reference sections of included studies were identified and manually added to the analysis. After initial abstract screening, 135 studies were eligible for full-text screening, with 76 excluded for the following reasons: no non-autistic control comparison (n = 12); not peer reviewed (n = 8), duplicated data (n = 10), not relevant (no MRS, MRSI used; n = 28), drug clinical trials with no recorded baseline (n = 4), the authors did not reply to requests for raw data and there were no other means of extracting raw data (n = 14). This is summarised in the PRIMSA chart (Figure 1).



**Figure 1. PRISMA flow diagram of study selection.**

Papers were included if they met all of the following criteria: (1) use of in-vivo MRS to measure metabolites in the brain; (2) investigation of brain metabolite concentrations in autism and non-autism groups (3) MRS measured from a specified brain region and (4) the study was published in a peer-reviewed journal and was written in English or translated to English. Reasons for exclusion; duplicate removal (n = 846), records excluded as not relevant after abstract screening (n = 2002), no case/control comparison (n = 12), no peer review (n = 8), duplicated data (n = 10), not using MRS (n = 28), drug trails with no reported baseline (n = 4), no access to data after emailing (n = 14).

## 2.2 Study selection and data extraction table

Relevant abstracts were independently assessed by two investigators (NAP and AT). This was followed by full text screening, with reasons for exclusion documented. Data were then extracted in 4 domains 1) brain metabolite levels in non-autism and autism cohorts. We note that although we chose non-autistic and not neurotypical (NT) to define the control group, as some control participants may simply be non-autistic and less care was taken to exclude other neurodevelopmental conditions 2) study cohort characteristics (i.e., sample size, age, gender, IQ); 3) reported MRS acquisition parameters (e.g., MRS sequence, MRI scanner static field strength, echo time (TE)/ repetition time (TR), voxel size, averages, processing software, pre- and post- processing parameters); 4) bibliometric data (e.g., authors, year of publication, and type of publication). Key correlative findings were also noted for each study for systematic review. Data extraction table can be found in supplementary table 1, final data table used for analysis is available at [https://osf.io/hxw3c/?view\\_only=5d2d683b027541879884485f138f5ddb](https://osf.io/hxw3c/?view_only=5d2d683b027541879884485f138f5ddb).

Concentrations of GABA, Glx, Glu, Cr, tCho, NAA and ml were recorded from reported mean and standard deviation (SD) values in autism and non-autism groups. If raw data values were not reported, lead authors were emailed, and where this was not successful, imputation methods were used in accordance with recommendations in the Cochrane meta-analysis handbook (Higgins & Green, 2002). As an example of the latter, if the data was displayed in a figure only, WebPlotDigitizer (Rohatgi, 2022) was used to extract data manually.

Studies vary in their metabolite quantification method: articles can report concentrations of neurometabolites as a ratio of a stable reference metabolite measured within the same voxel, this usually being creatine. Alternatively, estimated neurometabolite concentrations in institutional units (i.u) are reported, whereby metabolite concentrations are calculated in reference to voxel water signal. As brain tissue types differ significantly in water composition, researchers should optimally control for differences in tissue composition between voxels and between groups by applying tissue correction. At the most basic level, metabolite concentrations are scaled according to the assumption that metabolite concentrations in CSF are negligible (Gasparovic et al. 2006; Harris et al., 2015). Additional corrections include ‘alpha correction’ of GABA concentrations, based on the assumption that GABA concentration is two times greater in grey matter (GM) compared to white matter (WM), as well as accounting for differences in voxel T1/T2 tissue relaxation (Jensen et al., 2005, Harris et al., 2015). We record if such corrections were performed for the relevant studies. Where studies reported estimated metabolite concentrations in i.u (water referenced) and creatine ratios, both were recorded, however as overall a greater number of studies reported concentrations in i.u, for the main analysis i.u was used in preference.

Across studies, specific metabolites were reported in various ways depending on the spectral separation achieved at each field strength, post-hoc processing corrections and assumptions made by individual researchers. For example, creatine was mostly reported as total creatine (tCr), this being the sum of creatine and its phosphorylated form phosphocreatine, as their peaks are hard to separate at 1.5/3 Tesla field strengths (Cr + pCr; N = 23). 4 articles reported creatine concentrations alone (Cr; N = 4). NAA was reported as

total NAA; the sum of the overlapping peaks N-acetyl-aspartate (NAA) and N-acetyl-aspartyl-glutamate (NAAG; N = 3), or NAA alone (N = 32). Note if it was not stated if NAA signal was composed of both NAA and NAAG, it was assumed researchers reported NAA alone (N=15). Finally, GABA was reported as GABA+ (GABA plus macromolecule contribution; N = 18), or as estimates of GABA concentrations alone (through using macromolecule suppression techniques/macromolecule basis set modelling; N = 8). Choline was always reported as a sum of the choline containing compounds; compounds phosphocholine, phosphatidylcholine and glycerophosphocholine (tCho; N = 27). Reports of GABA and GABA+ were grouped under “GABA”, reports of tNAA and NAA were grouped under “NAA” and reports of Cr + pCr and tCr were grouped under “creatine” for ease of analysis and to optimise inclusion for statistical power. Glx and Glu data was analysed separately.

### **2.3 Quality assessment**

#### AXIS and MRS-Q

A modified appraisal tool for cross-sectional studies, specifically developed for MRS data, was used to assess cross sectional study design and reporting using 19 domains (AXIS; Downes et al., 2016; Peek et al., 2020). For each study, the percentage of AXIS domains satisfied was assessed. Domain 9 “were the risk factors and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously” of AXIS was discounted. Instead, the consensus quality appraisal tool ‘MRS-Q’ was used to comprehensively assess if MRS acquisition parameters satisfied consensus standards, and specifically, if reporting of MRS parameters was sufficient (A. Lin et al., 2021; Peek et al., 2020).

Studies that reported sufficient spectroscopy parameters to satisfy MRS-Q consensus standards for data acquisition were categorised as ‘high quality’ data. Studies that reported insufficient spectroscopy parameters to satisfy the consensus standards were classified as ‘low quality’ data, and studies with not enough information to make judgement were considered ‘unsure’. Importantly, this classification may fail to reflect the true quality of the study itself but reflects that reporting was poor. We do however analyse data with and without ‘low-quality’ studies to assess the impact of reported data quality on outcome. Two investigators (NAP and AT) independently assessed the quality of each study for both criteria.

## 2.4 Publication bias

Effect sizes were aggregated for each metabolite within each study (across the brain regions studied) using the ‘aggregate’ function from the *metafor* package in R (Viechtbauer, 2010). This attempted to increase the homogeneity of effect sizes and as such usefulness of publication bias tests, as well as preventing non-independence effects (due to multiple brain regions reported from the same participants in each study).

Egger’s regression tests (Bowden et al., 2015; Egger et al., 1997; Pasanta et al., 2023) were performed on aggregated metabolite data using the *meta* package function ‘metabias’ (Balduzzi et al., 2019). For Egger’s regression test (Egger et al., 1997; Sterne & Egger, 2001) a funnel plot was generated for each metabolite with effect size per study (across all brain regions) on the x axis and the standard error of the effect size on the y axis. In theory,

without publication bias, effect sizes should follow a symmetrical distribution around 0 (L. Lin & Chu, 2018). Egger's regression quantifies if effect sizes are asymmetrical and so may have a publication bias.

The trim-and-fill method is a non-parametric test used to correct for asymmetry of effect sizes due to publication bias (Duval & Tweedie, 2000; Pasanta et al., 2023). First, the studies driving asymmetry (extreme effect sizes) are identified and 'trimmed', and then a new estimate for the true centre (true effect size) of the distribution is found. This distribution is then 'filled' as the trimmed studies are added back, along with a prediction of the studies not reported due to publication bias (Balduzzi et al., 2019; Duval & Tweedie, 2000). For the trim-and-fill, a random-effects model using the Knapp and Hartung method was used on aggregated data using the *metabias* package '*funnel.meta*' and '*trimfill*' function (Balduzzi et al., 2019; Nakagawa et al., 2022; Pasanta et al., 2023). Data distribution was visualized on funnel plots (Begg & Mazumdar, 1994; Sterne & Egger, 2001).

## 2.5 Data analysis

Data-analysis was performed on extracted data using the *Meta-Essentials* tool in R (Suurmond et al., 2017). Standardized mean differences (Hedge's G) and 95% confidence intervals were calculated from mean metabolite concentrations and standard deviation in the autism and non-autistic group (per study per metabolite), allowing us to compare data reported in different units (creatine scaled and i.u). Subgroup and meta-regression analyses (*robumeta* package) were used to explore if effect sizes were associated with demographic and methodological parameters. Note for all metabolites, mean non-autistic concentration

was subtracted from mean autistic concentration and as such standardised mean differences (effect size) are expressed with respect to autism. For example, a negative effect size indicates lower concentrations of said metabolite in autism while a positive effect size indicates higher concentrations in autism compared to non-autistic controls.

Most studies acquired multiple voxels (brain regions) and metabolites from the same cohort, and so several effect size estimates were obtained from a single study e.g. NAA, Creatine and Glx from a frontal and temporal voxel per participant. This can lead to statistical dependency between measures, potentially skewing the calculation of group effect size parameters (Borenstein et al., 2021). To control for this, we use robust variance estimation (RVE) for group effect size calculations (Hedge's  $g$ ), using the *robumeta* package in R (Fisher & Tipton, 2015; Pasanta et al., 2023; Pustejovsky & Tipton, 2022). RVE assumes that effect sizes from the same study are correlated, and since the exact correlation is unknown, RVE estimates this dependence structure. A conservative correlation coefficient of 0.7 was used for all RVE calculations in accordance with (Pasanta et al., 2023; Rosenthal, 1986). Further adjustments of RVE estimates were made to control for small samples sizes (Tipton, 2015). Furthermore, when degrees of freedom were smaller than 4, a significance threshold of  $p < 0.01$  was used (instead of  $p < 0.05$ ) to assess significance of Hedge's  $g$ , as type 1 errors are more likely in this case (Tanner-Smith et al., 2016).

Data heterogeneity was evaluated using  $I^2$  (Higgins et al., 2002) and  $\tau^2$  (Viechtbauer, 2005, 2010).  $I^2$  statistic is an estimate of the proportion of variance in effect sizes (between studies) that reflects true heterogeneity, ranging from 0 to 100. High  $I^2$  ( $>70$ ), suggests there may be real differences between study outcomes not as a result of sampling error



(instead due to real differences in study methods/cohort characteristics) and so sub-group analysis maybe more appropriate (Borenstein et al., 2021; Hak et al., 2016; Higgins & Green, 2002).  $\tau^2$  squared is the estimated variance of true effect sizes, in other words measuring the amount of true heterogeneity. If the confidence interval for  $\tau^2$  does not contain 0, this suggests that between study heterogeneity exists and again subgrouping may be appropriate (Viechtbauer, 2005, 2010).

## 2.6 Subgrouping

Our main aim was to identify general differences in GABA, Glu, Glx, creatine, choline, GSH, NAA and ml concentrations between autistic and non-autistic cohorts and explore if this varies with age, brain region and study methodology. As such, we sub-grouped the data based on metabolite of interest, age group of cohorts (adults  $\geq 18$  years of age, children  $< 18$  years of age) and brain region. We then grouped data by demographic factors (sex, IQ) and acquisition parameters (scanner strength, sedation, pulse sequence etc.). We additionally looked at sex ratios in each study, testing both a 1:3 female:male ratio (as diagnostic bias in autism is currently 1:3 (Loomes et al., 2017), and 1:1 ratio for sex-balanced studies). The forestplot function in the *forestplot* package in R was used to plot RVE estimates of Hedges'  $g$  per subgroup. We also used the *cor* function in *r* to estimate the Spearman's rho correlation coefficient between voxel size/number of transients and study effect size.

To optimise study numbers per brain region, we grouped data into broader regions (brain region grouping 1; supplementary table 2) using previously defined and functionally relevant brain regions groupings (Pasanta et al., 2023). We use a secondary brain region grouping

(brain region grouping 2; supplementary table 3) to explore significant findings with more spatial resolution, however study number per group was considerably reduced (e.g.  $n = 2-4$ ) and as such brain grouping 1 was considered more appropriate and preferred for main interpretations.

Finally, as we aimed to explore whether there was an association between study outcome and quality of MRS data, we repeated the main analysis including only high-quality studies as defined by MRS-Q.

### **3. Results**

#### **3.1 MRS-Q**

Most studies ( $n = 46/59$ , 88%) reported methods that satisfied the MRS-Q criteria of best practice and so were deemed to be of high quality (Figure 2). Ten studies (17%) were assessed as low quality; this was mostly due to small voxel sizes and too few transients. All low-quality studies used conventional (non-editing) sequences (so were not observing GABA; Bernardi et al., 2011; Goji et al., 2017; Harada et al., 2011; Hardan et al., 2008, 2016; Otsuka et al., 1999), while 2 also used edited MRS sequences in combination with conventional MRS (Goji et al., 2017; Harada et al., 2011). Note for (Goji et al., 2017) and (Harada et al., 2011) only conventional (non-editing) MRS data was classified as low quality, edited MRS acquisition parameters were assessed separately and classified as high quality. For the high quality studies, 27 studies used conventional (non-editing) MRS sequences (L. Lin & Chu, 2018). 12 used editing MRS sequences (Bernardino et al., 2022; DeMayo et al.,

2021; Drenthen et al., 2016; Fung et al., 2021; Gaetz et al., 2014; Kirkovski et al., 2018; Pierce et al., 2021; Port et al., 2019; Puts et al., 2017; Rojas et al., 2014; Sapey-Triomphe et al., 2021; Umesawa et al., 2020). 8 studies used both editing and conventional (non-editing) sequences (Brix et al., 2015; Carvalho Pereira et al., 2018; Cochran et al., 2015; He et al., 2021; Ito et al., 2017; Maier et al., 2022; Moxon-Emre et al., 2021; Pierce et al., 2021; Wood-Downie et al., 2021). Three studies were classed as unsure due to missing information regarding MRS parameters (Drenthen et al., 2016; Kubas et al., 2012; Tebartz van Elst et al., 2014).

The poorest performing domains for MRS-Q were: reporting of scanner drift (only 10% of studies), reporting of number of MRS data points (30% of studies) and reporting of power calculations for sample size/justification of chosen sample size (only 16% of studies).

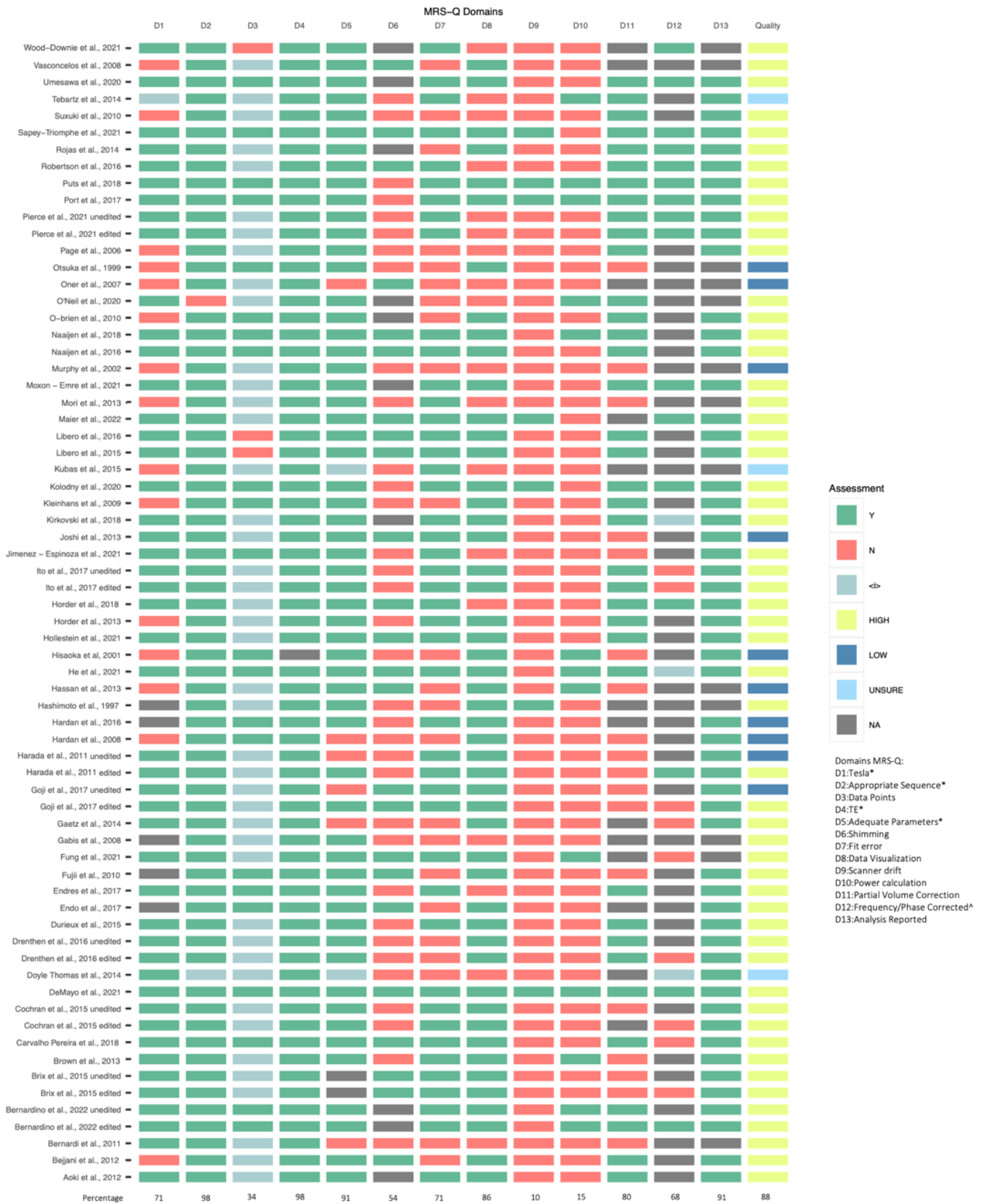
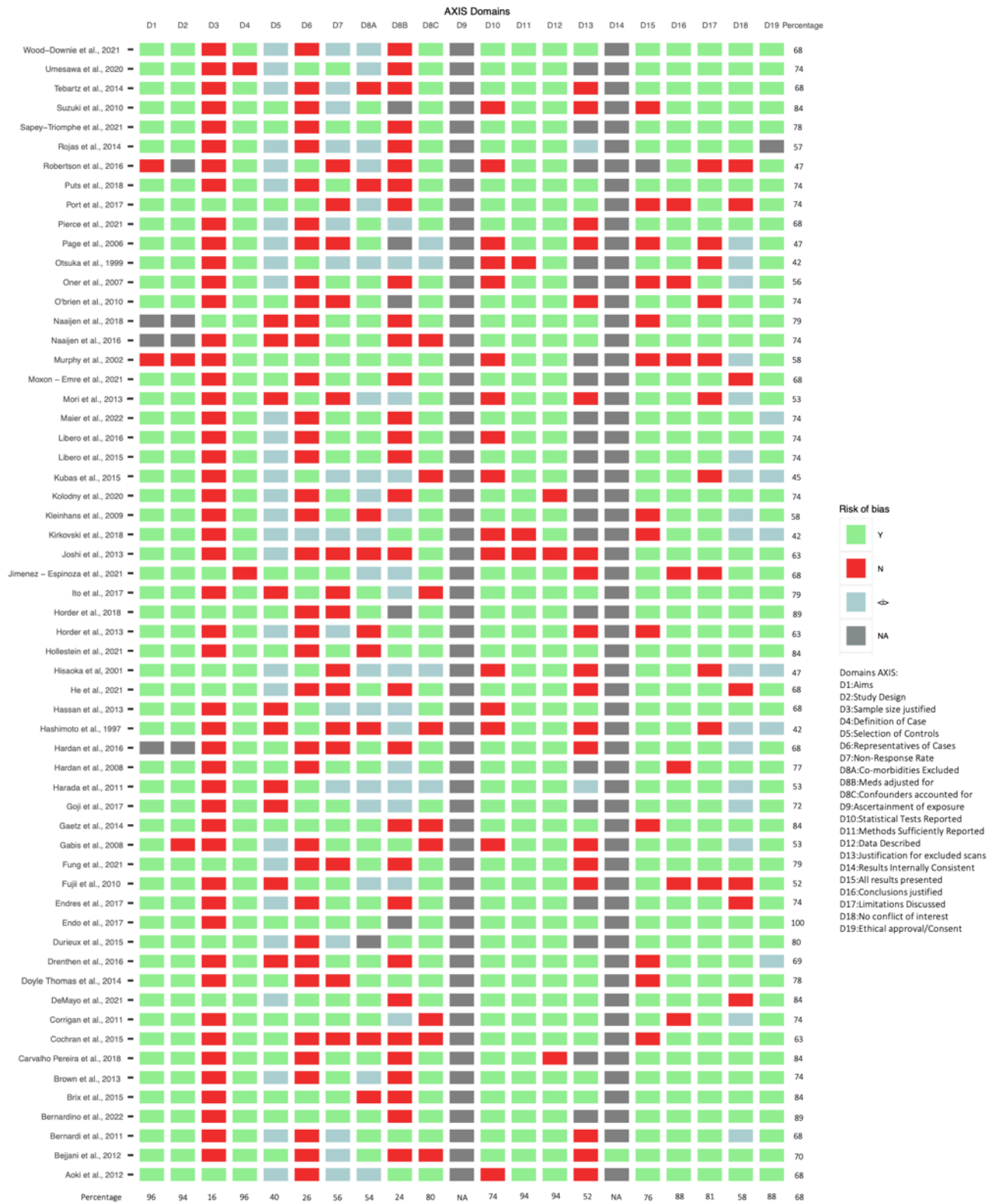


Figure 2: MRS quality as assessed by MRS-Q.

MRS-Q assessment of MRS studies, domain titles are shown in the key. \* Indicates the domain was used for quality assessment of MRS acquisition and reporting. ^ indicates that this domain is only applicable to edited MRS.

### 3.2 AXIS

While on average the included studies satisfied 68% of the AXIS criteria (Figure 3), the quality of the studies varied. Eight studies satisfied over 80% of the criteria (Bernardino et al., 2022; Brix et al., 2015; DeMayo et al., 2021; Durieux et al., 2016; Gaetz et al., 2014; Hollestein et al., 2021; Horder et al., 2018; Suzuki et al., 2010), while 10 studies satisfied 50% or less of the criteria only (Gabis et al., 2008; Harada et al., 2011; Hashimoto et al., 1997; Hisaoka et al., 2001; Kirkovski et al., 2018; Kubas et al., 2012; Oner et al., 2007; Otsuka et al., 1999; Page et al., 2006; Robertson et al., 2016). 96% of studies clearly defined an aim, the 'case' population and included sufficient methods for replication, however, only 26% of studies selected a representative control population from the same local population as case populations, only 56% of studies categorized non responders and 52% justified excluded scans. Studies also varied in controlling for confounding factors; 54% explicitly stated that co-occurring conditions were excluded, while only 24% of studies excluded or adjusted for medication use. Other confounders such as sedation, IQ, and age were mostly controlled, mostly post-hoc statically (80% of studies). As we extract raw data, these post-hoc statistical controls do not stand in our meta-analysis. Most studies (around 70%) presented all results and justified their conclusions.



**Figure 3: Study design assessed by AXIS.**

A modified appraisal tool for cross-sectional studies (AXIS), specifically developed for MRS, was used to assess the quality of MRS study design and reporting.

### 3.3 Publication bias assessment

#### Egger's regression test

No significant funnel plot asymmetry was found for NAA, choline, creatine, Glx, Glu and ml (Table 1). Significant asymmetry ( $p < 0.05$ ) was identified for studies observing GABA. Funnel plots suggest an absence of positive effect sizes. Note all GABA studies with a large negative effect size, and so driving this asymmetry, reported creatine-scaled metabolite concentrations (Gaetz et al., 2014; Kubas et al., 2012; Port et al., 2019). For GSH, too few studies had been performed ( $n = 3$ ) for appropriate bias assessment.

**Table 2. Egger's Regression test for publication bias (funnel plot asymmetry).**

| Metabolite | Bias    | Se     | Intercept | Se     | t     | p      | df | Result  |
|------------|---------|--------|-----------|--------|-------|--------|----|---------|
| NAA        | -2.0934 | 1.1143 | 0.3898    | 0.3474 | -1.88 | 0.0692 | 33 | No bias |
| choline    | 0.8477  | 0.8527 | -0.2251   | 0.2625 | 0.99  | 0.3297 | 25 | No bias |
| creatine   | -1.4758 | 1.5770 | 0.3576    | 0.5106 | -0.94 | 0.3583 | 25 | No Bias |
| Glx        | -0.9139 | 1.3760 | 0.2388    | 0.4202 | -0.64 | 0.5301 | 27 | No Bias |

|      |         |        |         |        |       |        |    |         |
|------|---------|--------|---------|--------|-------|--------|----|---------|
| Glu  | 0.4854  | 1.4268 | -0.0136 | 0.3854 | 0.34  | 0.7401 | 11 | No Bias |
| GABA | -3.1769 | 1.2545 | 0.6360  | 0.3785 | -2.53 | 0.0183 | 24 | Bias    |
| GSH  | NA      | NA     | NA      | NA     | NA    | NA     | NA | NA      |
| ml   | 0.4494  | 1.4071 | -0.2171 | 0.5464 | 0.32  | 0.7531 | 18 | No Bias |

Note: significance indicates an asymmetrical distribution of effect sizes on funnel plots.

### Trim-and-fill

Three studies were ‘trimmed and filled’ for NAA (Gabis et al., 2008; Hardan et al., 2016; Kubas et al., 2012) and 2 for ml (Brown et al., 2013; Gabis et al., 2008), with the resulting estimated effect sizes reported in Table 3.

All data demonstrated high heterogeneity, with  $I^2$  calculated at 70-80% even when data was split by metabolite (Figure 3). This suggests that more than half of the variation in estimated effect sizes is due to true differences in effect size between studies, likely because of the differences in brain region, age etc. It has been advised in such cases publication bias analysis may not be appropriate (Nakagawa et al., 2022; van Aert et al., 2016). As such, identified ‘trimmed’ studies remain in the analysis, while publication bias results were considered during interpretation.



**Table 3. Trim-and-Fill correction for publication bias.**

| Metabolite | Filled              | Adjusted Estimate<br>effect size | 95%CI            | T value | p      | tau <sup>2</sup> |
|------------|---------------------|----------------------------------|------------------|---------|--------|------------------|
| NAA        | Kubas et al., 2012  |                                  |                  |         |        |                  |
|            | Gabis et al., 2008  | -0.1771                          | -0.5230; 0.1689  | -1.04   | 0.3064 | 0.7710           |
|            | Hardan et al., 2016 |                                  |                  |         |        |                  |
| choline    | None                | 0.0310                           | -0.1137; 0.1756  | 0.44    | 0.6636 | 0.0115           |
| creatine   | None                | -0.1369                          | -0.4039; 0.1301  | -1.05   | 0.3015 | 0.2630           |
| Glx        | None                | -0.0417                          | -0.3094; 0.2260  | -0.32   | 0.7519 | 0.2868           |
| Glu        | None                | 0.1086                           | -0.2179; 0.4351  | 0.72    | 0.4825 | 0.1230           |
| GABA       | None                | -0.4124                          | -0.7505; -0.0743 | -2.51   | 0.0188 | 0.4351           |
| GSH        | None                | 0.0962                           | -1.3191; 1.5115  | 0.29    | 0.7975 | 0.2318           |
| MI         | Brown et al., 2012  |                                  |                  |         |        |                  |
|            | Gabis et al., 2008  | -0.1914                          | -0.6107; 0.2278  | -0.95   | 0.3531 | 0.5973           |

### 3.4 Study characteristics

All study characteristics are shown in Table 4.

**Table 4. Study characteristics**

| Author                             | Metabolite studied                | Voxel placement                   | Voxel size                      | Same size (autism: non-autism) /<br>Age (autism: non-autism) /<br>% females (autism group) | Scanner strength /<br>Sequence /<br>TR:TE (ms) /<br>Number of transients / Quantification method | participants on<br>psychoactive medication /<br>Use of sedation /<br>Diagnosis criteria /<br>Diagnosis /<br>IQ > 70? | MRS-Q<br>quality |
|------------------------------------|-----------------------------------|-----------------------------------|---------------------------------|--|--|--|------------------|
| (Aoki et al., 2012)                | NAA, tCho, Cr, Glx,<br>ml         | Prefrontal                        | 20 x 20 x 20 cm3                | 24:25 / 29.5:29.4 / 0%   | 3T / STEAM / 15:3000 ms / 128 / i.u  | Y / N / DSM - 4 / ASD / Y  | HIGH             |
| (Bejjani et al., 2012)             | tNAA, tCho, tCr, Glx,<br>ml       | Limbic                            | 3.6 cm3                         | 8:10 / 11.2 : 13.2 / 12.5%   | 1.5T / PRESS / 25: 1500 ms / <i> / i.u   | Y / Y / DSM -4 & ADOS &<br>ADI-R / ASD / Y   | HIGH             |
| (Bernardi et al., 2011)            | NAA, tCho, tCr, Glx,<br>ml        | Limbic, Parietal and<br>temporal  | 0.56 cm3                        | 12:14 / 29.2:29.7 / 16%  | 3T / PRESS / 30:2000ms / <i> / i.u   | N / N / DSM-4 & ADI-R &<br>ADOS-G / ASD / Y  | LOW              |
| (Bernardino et al., 2022)          | GABA+                             | Frontal                           | 27ml                            | 14:15 / 20.6:22.9 / 0%   | 3T / MEGA-PRESS / 35:2000ms / 196 / i.u  | Y / N / DSM-4 & ADI-R &<br>ADOS-G / ASD / Y  | HIGH             |
| (Brix et al., 2015)                | NAA, tCho, tCr, Glx,<br>GABA+, ml | Limbic                            | 27ml                            | 14:24 or 21 / 10.2:10.2 / 0%   | 3T / PRESS & MEGA-PRESS / 35:1500ms and<br>68:15000ms / 128 / i.u                                | Y / N / DSM-5 / ASD / N  | HIGH             |
| (Brown et al., 2013)               | tNAA, Glx, tCr, ml,<br>tCho       | Temporal                          | 25 x 14 x 20<br>mm3             | 13:15 / 37: 41 / 31%   | 3T / PRESS / 20:2000ms / 192 / i.u   | Y / N / DSM-4 / ASD / Y  | HIGH             |
| (Carvalho Pereira et al.,<br>2018) | tNAA, tCr, Glx,<br>GABA+          | Prefrontal                        | 27ml                            | 21:14 / 13:13 / <i>  | 3T / MEGA-PRESS / 68:1500ms / 392 / i.u  | Y / N / ADI-R & ADOS &<br>DSM-4 / ASD / Y  | HIGH             |
| (Cochran et al., 2015)             | Glu, GABA                         | Limbic areas                      | 20x20x20mm3<br>30x30x20mm3      | 13:14 / 15:14.7 / 0%   | 3T / PRESS & MEGA-PRESS / 28:2000ms or<br>68:2000ms / 230 or 128 / i.u & creatine                | Y / N / DSM-4 & ADOS /<br>ASD / Y  | HIGH             |
| (DeMayo et al., 2021)              | GABA+                             | Parietal                          | 27mm3                           | 24:35 / 8.77:8.46 / 4%   | 3T / MEGA-PRESS / 68:1800ms / 256 / i.u  | Y / N / ADOS-2 / ASD / Y   | HIGH             |
| (Doyle-Thomas et al., 2014)        | Glx                               | Limbic                            | 27ml                            | 20:16 / 9.8:11.1 / 0%  | 3T / <i> / 30:2000ms / <i> / creatine  | N / N / ADI-R % ADOS &<br>DSM-4 / ASD / Y  | UNSURE           |
| (Drenthen et al., 2016)            | Glu, GABA+                        | Occipital                         | 3x3x3 cm3                       | 15:15 / 14.9:14.7 / 0%   | 3T / PRESS & MEGA-PRESS / 35:2000ms or<br>68:2000ms / 128 or 320 / creatine                      | Y / N / DSM-4 / ASD / Y  | HIGH             |
| (Durieux et al., 2016)             | GSH                               | Prefrontal, limbic                | 20x20x15mm2                     | 21:29 / 31.8:26.9 / 0%   | 3T / PRESS / 30:3000ms / 128 / i.u   | N / N / ICD-10 / ASD / Y   | HIGH             |
| (Endo et al., 2007)                | NAA, tCho, Cr, GSH,<br>ml         | Limbic, Prefrontal,<br>Cerebellum | 20x20x20 cm3                    | 38:16 / 13.1:11.5 / 50%  | 1.5T / PRESS / 35:2000ms / 128 / creatine  | N / N / DSM-4 / ASD / N  | HIGH             |
| (Endres et al., 2017)              | NAA, tCho, Cr, GSH,<br>ml         | Limbic, Prefrontal                | 20x30x30mm3                     | 24:18 / 25-75 / 50%  | 3T / PRESS / 30:1500ms / 256 / i.u   | N / N / ICD-10 & ADI-R /<br>ASD / Y  | HIGH             |
| (Fujii et al., 2010)               | NAA, tCho, tCr                    | Prefrontal, Limbic                | 15x15x15mm3                     | 20: 18 / 7.5:7.8 / 15%   | 1.5T / PRESS / 135:1300ms / 256 / i.u  | <i> / N / DSM-4 / AD / N   | HIGH             |
| (Fung et al., 2021)                | NAA, tCr, Glx, GABA               | Limbic, Prefrontal                | 15cm3                           | 28:29 / 26.6:27.7 / 39%  | 3T / MEGA-SPECIAL / 80:2000ms / <i> / creatine<br>& i.u  | N / N / DSM-5 & ADOS-2 /<br>ASD / Y  | HIGH             |
| (Gabis et al., 2008)               | NAA, tCho, ml, tCr                | Limbic, Cerebellum                | 8ml                             | 13:8 / 10:11 / 23%   | 1.5T / PRESS / 40:2000ms / <i> / creatine  | N / N / DSM-4 / ASD / Y  | HIGH             |
| (Gaetz et al., 2014)               | GABA+                             | Frontal, Temporal,<br>Occipital   | 27mm3                           | 17:15 / 11.5:12.7 / 5%   | 3T / MEGA-PRESS / 68:1500ms / 128 / creatine   | Y / N / DSM-4 / ASD / N  | HIGH             |
| (Goji et al., 2017)                | NAA, tCho, tCr, Glx,<br>GABA      | Limbic, Occipital                 | 15x20x20 mm3<br>1.5 x 2 x 3 cm3 | 34: 19 / 5.2:5.6 / 18%   | 3T / STEAM&MEGA-PRESS / 15:5000ms or<br>68:2500ms / 46 & 256 / i.u                               | <i> / <i> / DSM-4 / AS / <i><br>Unedited:<br>LOW<br>Edited:<br>HIGH  |                  |
| (Harada et al., 2011)              | NAA, Cr, Glu,<br>GABA+, ml        | Frontal, Limbic                   | 27mm3                           | 12:10 / 5.2:5.9 / <i>  | 3T / MEGA-PRESS & STEAM / 15:5000ms or<br>68:2500ms / 256 & 48 / i.u                             | <i> / <i> / DSM-4 / ASD / <i><br>Unedited:<br>LOW<br>Edited:<br>HIGH   |                  |
| (Hardan et al., 2008)              | NAA, tCho, tCr, Glx,<br>ml        | Limbic                            | 15x20x20mm3                     | 18 or 12: 16 / 11.9:11.6 / 0%  | 1.5T / STEAM / 20:1600ms / 4 / i.u   | <i> / N / ADI-R & ADOS /<br>AD / Y   | LOW              |

|  |                           |   |  |   |  |   |
|--|---------------------------|---|--|---|--|---|
| (Hardan et al., 2016)                          | NAA, tCho, ml             | Frontal, Limbic, Parietal               | 15x15x20mm <sup>3</sup>                            | 17:17 / 12.5:11.6 / <i>                           | 1.5T / STEAM / 20:1600ms / 4 / i.u   | Y / N / ADI-R & ADOS / ASD LOW / Y            |
| (Hashimoto et al., 1997)                       | NAA, tCho                 | Parietal                                | 8-27ml   | 28:25 / 5.8:6.4 / 28%                             | 1.5T / STEAM / 270:1500ms / 200-500 / creatine <i> / <i>                     | DSM-III / AD / <l. HIGH                       |
| (Hassan et al., 2013)                          | Glu                       | Limbic, Cerebellum, Frontal             | 27ml   | 10:10 / 11.4:11.3 / 40%                           | 1.5T / PRESS / 30:500ms / 3 / i.u  | N / Y / DSM-4 / AD / Y LOW                    |
| (He et al., 2021)                              | GABA+, Glx, GSH           | Parietal, Limbic                        | 27ml   | 44:62 / 10.3: 9.69 / 4%                           | 3T / MEGA-PRESS & HERMES / 80:2000ms / 320 / i.u                             | Y / N / DSM-4 & ADOS-2 & ADI-R / ASD / Y HIGH |
| (Hisaoaka et al., 2001)                        | NAA                       | Frontal, Temporal, Parietal, Brain stem | 20x20x20 cm <sup>3</sup>                           | 55:51 / 5.8: 6.2 / 15%                            | 1.5T / PRESS / 135:1300ms / 256 / i.u  | <i> / N / ICD-10 / AD / <i> LOW               |
| (Hollestein et al., 2021)                      | Glu                       | Limbic                                  | 20x20x20 cm <sup>3</sup>                           | 24:15 / 11.38:10.7 / 29%                          | 3T / PRESS / 30:3000ms / 96 / i.u  | N / N / DSM-4 & ADI-R / ASD / ASD / Y HIGH    |
| (Horder et al., 2013)h                         | NAA, tCho, tCr, Glx       | Limbic, Prefrontal, Parietal            | 20x20x20cm <sup>3</sup>                            | 13:13 / 27:34 / 3%                                | 1.5T / PRESS / 30:3000ms / 256 / i.u   | N / N / ICD-10 & ADI-R / ASD / Y HIGH         |
| (Horder et al., 2018)                          | NAA, tCho                 | Prefrontal, Limbic                      | 20x30x20cm <sup>3</sup>                            | 25:36 / 30.98:28.91 / <i>                         | 3T / PRESS / 68:2000ms / 256 / i.u   | N / N / ICD-10 & ADOS-G / ASD / Y HIGH        |
| (Ito et al., 2017)                             | NAA, tCho, GABA+, ml, tCr | Limbic, Cerebellum                      | 15x20x30mm <sup>3</sup>                            | 112 or 114: 65 or 45/ 6.4:6.7 or 5:2 / 30% or 35% | 3T / STEAM & MEGA-PRESS / 15:5000ms / creatine                               | <i> / Y / DSM-4 / ASD / N HIGH                |
| (Jiménez-Espinoza et al., 2021)                | NAA, tCho, Cr, ml         | Limbic                                  | 20x20x20cm <sup>3</sup>                            | 27:44 / 20.58:23.19 / 11%                         | 3T / PRESS / 23:1070ms / i.u   | <i> / N / ADI-R & ADOS / ASD / <i> UNSURE     |
| (Joshi et al., 2013)                           | Glu                       | Limbic, Temporal                        | 3.73cm <sup>3</sup>                                | 7:7 / 14:<i> / 0%                                 | 4T / JPRESS / 250:2000ms / i.u   | Y / N / DSM-4 / AD / Y LOW                    |
| (Kirkovski et al., 2018)                       | GABA+                     | Temporal                                | 27ml   | 12:14 / 28:31 / 30%                               | 3T / MEGA-PRESS / 68:1500ms / i.u  | N / N / ADI-R / ASD / <i> HIGH                |
| Kleinhans et al., 2009(Kleinhans et al., 2009) | NAA, tCho, tCr, ml        | Limbic                                  | 15x15x15 mm <sup>3</sup>                           | 19:23 / 23:23 / 10%                               | 1.5T / PRESS / 30:2000ms / i.u   | <i> / N / DSM-4 & ADI-R & ADOS / ASD / Y HIGH |
| (Kolodny et al., 2020)                         | Glx                       | Occipital                               | 27ml   | 29:38 / 22:23 / 72%                               | 3T / PRESS / 68:2000ms / 320 / i.u   | Y / N / ADI-R & ADOS / ASD HIGH / Y           |
| (Kubas et al., 2012)                           | NAA, Glx, ml              | Limbic                                  | <i>  | 12:16 / 10:11 / 41%                               | 1.5T / PRESS / 36:1500ms / <i> / creatine                                    | <i> / N / ICD-10 / AD / <i> UNSURE            |
| (Liberio et al., 2015)                         | Glx, tCho, tNAA           | Limbic                                  | 20x27x20mm <sup>3</sup>                            | 19:18 / 27:24 / 21 %                              | 3T / PRESS / 80:2000ms / 128 / creatine                                      | Y / N / ADI-R & ADOS / ASD / <i> HIGH         |
| (Liberio et al., 2015)                         | Glx, tCho, tNAA           | Limbic                                  | 20x27x20 mm <sup>3</sup>                           | 20:20 / 26:24 / 25%                               | 3T / PRESS / 80:2000ms / 128 / i.u   | Y / N / ADI-R & ADOS / ASD HIGH / Y           |
| (Maier et al., 2022)                           | Glx, GABA                 | Prefrontal, Limbic, Prefrontal          | 20x20x15 mm <sup>3</sup>                           | 43: 43 / 11:34 / 29%                              | 3T / PRESS & MEGA-PRESS / 512 / i.u  | Y / N / ICD-10 & DSM-4 / ASD / Y HIGH         |
| (Mori et al., 2013)                            | NA, tCho, tCr             | Limbic, Prefrontal                      | 15x15x15mm <sup>3</sup>                            | 77:31 / 4:4 / 25%                                 | 1.5T / STEAM / 15:5000ms / <i> / i.u   | <i> / Y / DSM-4 / AD / N HIGH                 |
| (Moxon-Emre et al., 2021)                      | Glx, GABA+                | Prefrontal                              | 27ml   | 19:28 / 23: 23 / 36%                              | 3T / MEGA-PRESS/ 68:1500ms / 512 / i.u                                       | Y / N / DSM-4 & 5 / ASD / Y HIGH              |
| (Murphy et al., 2002)                          | NAA, tCho, tCr, ml        | Prefrontal, Parietal                    | 8.5ml  | 11:18 / 30:32 / 0%                                | 1.5T / PRESS / 136:2000ms / <i> / i.u  | N / N / ICD-10 / AS / Y LOW                   |
| (Naaijen et al., 2017)                         | Glu                       | Limbic                                  | 20x20x20mm <sup>3</sup>                            | 51:53 / 10:10 / 24%                               | 3T / PRESS / 30:3000ms / 96 / i.u  | Y / N / DSM-4 & 5 / ASD / Y HIGH              |
| (O'Brien et al., 2010)                         | tCho, tCr, ml, NAA        | Limbic                                  | 20x20x15mm <sup>3</sup>                            | 22:22 / 23:23 / 18%                               | 1.5T / PRESS / 35:3000ms / 160 / i.u   | N / N / ADI-R & ADOS / AS HIGH / Y            |
| (Oner et al., 2007)                            | NAA, tCho                 | Limbic, Prefrontal                      | 10x10x10mm <sup>3</sup>                            | 14:21 / 24:25 / 0%                                | 1.5T / PRESS / 270:1500ms / <i> / creatine                                   | Y / N / DSM-4 / AS / Y LOW                    |
| (Otsuka et al., 1999)                          | NAA, tCho, tCr            | Limbic, Cerebellum                      | 2x2x1.5 cm <sup>3</sup>                            | 27:10 / 2-18 years/ <i>                           | 1.5T / STEAM / 18: 5000ms / <i> / i.u  | <i> / N / ICD-10 / AD / <i> LOW               |
| (Page et al., 2006)                            | NAA, tCho, tCr, Glx, Mi   | Limbic, Parietal                        | 8ml  | 25 :21/ 11:9 / 25%                                | 1.5T / PRESS / 35:3000ms / 160 / i.u   | N / N / ICD-10 & ADI-R / ASD / Y HIGH         |
| (Page et al., 2006)                            | GABA+ Glx                 | Frontal, Parietal, Occipital            | 20x30x30mm <sup>3</sup><br>20x30x20mm <sup>3</sup> | 16:16 / 11.3:10/18%                               | 3T / MEGA-semi-LASER & semi-LASER / 35:2000ms or 68:2000ms / 128 or 64 / i.u | <i> / N / DSM-4 / ASD/ <i> HIGH               |
| (Port et al., 2019)                            | GABA+                     | Temporal                                | 27ml   | 30:27/21:27 / 2%                                  | 3T / MEGA-PRESS / 256 / creatine   | Y / N / DSM-4& ADOS / ASD / Y HIGH            |
| (Puts et al., 2017)                            | GABA+                     | Parietal                                | 27ml   | 37:35 / 10:10 / 21%                               | 3T / MEGA-PRESS / 256 / i.u  | Y / N / DSM-4 & ADOS-G / ASD / Y HIGH         |
| (Robertson et al., 2016)                       | GABA, Glx                 | Occipital                               | 2.5x2.5x3cm <sup>3</sup>                           | 20:21/29:29 / 20%                                 | 3T / MEGA-PRESS & PRESS/ 68:2000ms / 320/ i.u                                | Y / N / DSM-4 / ASD/ Y HIGH                   |
| (Rojas et al., 2014)                           | GABA+                     | Temporal                                | 30x30x40mm <sup>3</sup>                            | 17:17/ 11:14 / 70%                                | 3T / MEGA-PRESS / 60:2400ms / 512 / Creatine                                 | Y / N / DSM-4 / ASD/ Y HIGH                   |
| (Sapey-Triomphe et al., 2021)                  | GSH, GABA+, Glx           | Occipital, Frontal                      | 27ml   | 26:26 / 30:32 / 30:32 / 81%                       | 3T / HERMES / 80:2000ms / 320 / Creatine & i.u                               | Y / N / DSM-5 / ASD/ Y HIGH                   |
| (Suzuki et al., 2010)                          | NAA, tCho, Cr,            | Limbic, Cerebellum                      | 2x2x1.5cm <sup>3</sup>                             | 12:12/22:22 / 0%                                  | 1.5T / PRESS / 144:1500ms / 256 / i.u  | Y / N / DSM-4 & ICD-10 / ASD / Y HIGH         |
| (Tebartz van Elst et al., 2014)                | NAA, tCho, Glx, Cr, Glu   | Limbic, Cerebellum                      | 20x20x20cm <sup>3</sup>                            | 29:29/ 35:35 / 34%                                | 3T / PRESS/ 30:3000ms / 256 / i.u  | <i> / N / DSM-4 & ICD-10 / ASD / Y HIGH       |
| (Umesawa et al., 2020)                         | GABA+                     | Frontal                                 | 20x20x15mm <sup>3</sup>                            | 21:20/ 19:19 / 28%                                | 3T/ MEGA-PRESS / 70:2000ms / 512 / i.u                                       | Y / N / <i> / ASD/ Y HIGH                     |

(Wood-Downie et al., 2021) Cr, Glx, GABA+ Limbic, Parietal 25x35x10mm3 35:33/ 14:13 / 57% 3T / PRESS & MEGA-PRESS / 68:2000ms/ 526/ Y / N / ADI-R & ADOS / HIGH  
i.u ASD/ Y

Key: <i> = missing information.

### 3.5 Meta-analysis results

Note all group effect sizes reported below are in respect to autism, as such a negative effect size suggests lower metabolite concentrations in autism and a positive effect size suggests higher metabolite concentrations in autism.

#### General trends

We first considered the change in concentration across all metabolites, brain regions, age groups and other factors (e.g., pulse sequence; Figure 4). We found a non-significant negative effect size (Hedges'  $g = -0.16$ , 95% CI:  $-0.34 - 0.022$ ,  $I^2 = 82$ ,  $n = 59$ , 57.6 % children,  $p = 0.085$ ).

#### *Grouped by age group.*

Based on reported age range of cohorts, 25 studies reported adult only cohorts ( $\geq 18$  years) and 32 studies reported child cohorts ( $< 18$  years). Note three studies included cohorts with a minimum age of 16/17 years, although had population mean age greater than 18 years and so cohorts were classified as adult (Bernardino et al., 2022, mean age = 29 years, range:

17 – 30 years; Jiménez-Espinoza et al., 2021, mean age = 20 years, range: unknown; Moxon-Emre et al., 2021, mean age = 23 years, range: 16 – 34 year).

When data were grouped by cohort age (adults or children), across all metabolites and brain regions (Figure 4), we observed a significant negative group effect size for studies observing metabolites in **children** (Hedges'  $g = -0.35$ , 95% CI:  $-0.70 - -0.009$ ,  $I^2 = 88$ ,  $n = 34$ ,  $p < 0.05$ ). In **adults**, we observed a non-significant group effect size close to zero (Hedges'  $g = -0.03$ , 95% CI:  $-0.24 - 0.18$ ,  $I^2 = 72$ ,  $n = 27$ ,  $p = 0.76$ ).

#### *Grouped by brain region.*

30 studies took measurements from more than one brain region. The most reported brain area (brain region grouping 1) was the limbic regions (39 studies; 181 observations), followed by the prefrontal regions (14 studies, 58 observations), the parietal regions (12 studies, 47 observations), the frontal regions (10 studies, 22 observations), the cerebellum (8 studies, 31 observations), temporal regions (8 studies, 19 observations) and occipital regions (6 studies, 16 observations).

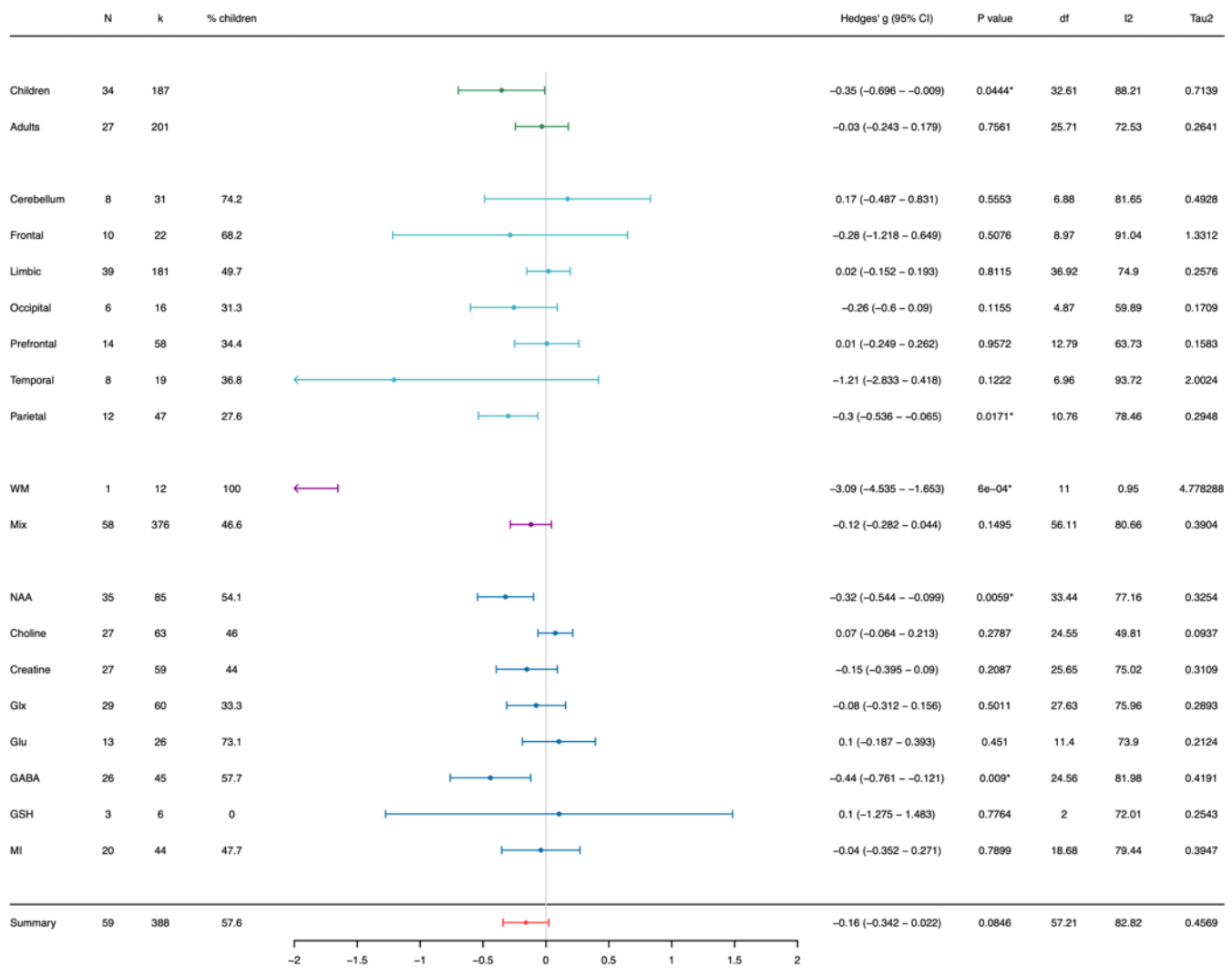
When data were grouped by brain region (grouping 1, across all metabolites and ages; Figure 4). We found a significant (negative) group effect size for studies observing metabolites in the parietal regions only (Hedges'  $g = -0.30$ , 95% CI:  $-0.54 - -0.065$ ,  $I^2 = 78$ ,  $n = 12$ ,  $p < 0.05$ ).

#### *Grouped by metabolite.*

43 studies observed more than one metabolite. NAA was the most studied metabolite (35 studies), followed by Glx (29 studies), choline (27 studies), creatine (27 studies), GABA (26 studies), ml (20 studies), Glu (13 studies) and GSH (3 studies). When the data were grouped by metabolite (across all brain regions and ages; Figure 4), group effect sizes for NAA and GABA were significant and negative (Hedges'  $g_{\text{NAA}} = -0.32$ , 95% CI: -0.54 - -0.099,  $I^2 = 77$ ,  $n = 35$ , 54.1 % children,  $p < 0.05$  and Hedges'  $g_{\text{GABA}} = -0.44$ , 95% CI: -0.76 - -0.12,  $I^2 = 82$ ,  $n = 26$ , 57.8 % children,  $p < 0.05$ ).

*Grouped by voxel tissue type.*

Data were also grouped by tissue type (WM/ GM/mixed) as one study reported metabolite concentrations in WM only. The majority (58) did not explicitly state the tissue type, and as such it was assumed that metabolite concentrations were assessed across the whole voxel (WM and GM), although voxel tissue composition likely varied greatly depending on brain region observed. The one study observing metabolite concentrations in WM only had a significant negative effect size (Hedges'  $g = -3.09$ , 95% CI: -4.54 - -1.65,  $I^2 = 0.95$ ,  $n = 1$ , 100% children,  $p < 0.01$ ; Figure 4). Across all brain regions, metabolites and ages, the other 58 studies that used mixed voxels (WM and GM) had a non-significant negative effect size (Hedges'  $g = -0.12$ , 95% CI: -0.28 - 0.28,  $I^2 = 56$ ,  $n = 58$ , 46.6% children,  $p = 0.15$ ; Figure 4).



**Figure 4.** Summary forest plot for data grouped by age, brain region, voxel tissue type, and metabolite of interest. N: number of studies, k: number of observations, % children: percentage of studies observing metabolite concentrations in children only cohorts, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism.

\*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.

To observe the effect of metabolite, brain region and age group simultaneously on effect sizes, data were split by:

1. **metabolite of interest and brain region (Figure 5).**
2. **metabolite of interest and age group (Figure 6).**
3. **metabolite of interest, age group, and brain region (Figure 7-13).**

Results for these subgroupings are reported per metabolite below. Note if  $n = 1$  for any subgroup explored, these results will be reported but excluded from interpretation.

## **NAA**

*By brain region* A significant negative group effect size was found for studies measuring NAA in **limbic** (Hedges'  $g = -0.3$ , 95% CI:  $-0.50 - -0.097$ ,  $I^2 = 64$ ,  $n = 28$ ,  $p < 0.05$ , 48.7% children; Figure 5) and **parietal** brain regions (Hedges'  $g = -0.34$ , 95% CI:  $-0.51 - -0.16$ ,  $I^2 = 0$ ,  $n = 6$ ,  $p < 0.05$ , 30% children; Figure 5). No significant effect sizes were observed for studies measuring NAA in the cerebellum ( $n = 7$ ; 75% children), frontal ( $n = 3$ ; 100% children), temporal ( $n = 2$ ; 75% children) or prefrontal regions ( $n = 12$ ; 46.7 % children).

To explore this in more spatial detail we grouped NAA data by brain region grouping 2, where limbic regions are differentiated. We observed a significant negative effect size in the ACC only (supplementary figure 2; Hedges'  $g = -0.22$ , 95% CI:  $-0.43 - -0.01$ ,  $I^2 = 41.5$ ,  $n = 14$ ,  $p < 0.05$ ). Note this region had considerably greater number of studies than all other limbic brain regions (basal ganglia, PCC, medial temporal lobe (hypothalamus-amygdala)).



*By age group* A significant negative group effect size was observed for studies measuring NAA in **children** (Hedges'  $g = -0.56$ , 95% CI:  $-1.00 - -0.12$ ,  $I^2 = 84$ ,  $n = 18$ ,  $p < 0.05$ ; Figure 6) but not in **adults** (Hedges'  $g = -0.09$ , 95% CI:  $-0.29 - 0.11$ ,  $I^2 = 55$ ,  $n = 18$ ,  $p = 0.36$ ; Figure 6).

*By age group and brain region* When studies measuring NAA were split by brain region and age group (Figure 7 – 13), a significant negative group effect size was observed for studies measuring **limbic** and **prefrontal** NAA in **children** (Hedges'  $g_{\text{limbic}} = -0.42$ , 95% CI:  $-0.80 - -0.041$ ,  $I^2 = 75$ ,  $n = 15$ ,  $p < 0.05$ ; Figure 9 and Hedges'  $g_{\text{prefrontal}} = -0.58$ , 95% CI:  $-1.10 - -0.067$ ,  $I^2 = 18$ ,  $n = 4$ ,  $p < 0.05$ ; Figure 10) but not in **adults** (Hedges'  $g_{\text{limbic}} = -0.16$ , 95% CI:  $-0.37 - 0.065$ ,  $I^2 = 39$ ,  $n = 14$ ,  $p = 0.15$ ; Figure 9 and Hedges'  $g_{\text{prefrontal}} = 0.11$ , 95% CI:  $-0.3 - 0.51$ ,  $I^2 = 59$ ,  $n = 8$ ,  $p = 0.56$ ; Figure 10).

For the cerebellum ( $n_{\text{children}} = 5$ ,  $n_{\text{adult}} = 2$ ; Figure 7), frontal ( $n_{\text{children}} = 3$ ,  $n_{\text{adult}} = 0$ ; Figure 8), parietal ( $n_{\text{children}} = 3$ ,  $n_{\text{adult}} = 4$ ; Figure 11), and temporal regions ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 1$ ; Figure 12) no significant effect size was observed for either age group for NAA.

To explore significant limbic findings in more detail, we grouped data by brain region grouping 2 (where limbic regions are differentiated) and age group. The ACC was the only limbic region with sufficient study numbers to explore age effects. We observed no significant effect size for ACC NAA concentrations in children (Hedges'  $g = -0.11$ , 95% CI:  $-0.40 - 0.17$ ,  $I^2 = 33$ ,  $n = 7$ ,  $p = 0.36$ ; supplementary figure 3) or adult (Hedges'  $g = -0.34$ , 95% CI:  $-0.71 - 0.03$ ,  $I^2 = 41$ ,  $n = 7$ ,  $p = 0.65$ ). ACC NAA findings are only significant when adult

and child data is collated (Supplementary figure 2; Hedges'  $g = -0.22$ , 95% CI:  $-0.43 - -0.01$ ,  $I^2 = 41.5$ ,  $n = 14$ ,  $p < 0.05$ ).

## **GABA**

*By brain region* We observed a significant negative group effect size for studies measuring GABA in the **limbic** regions (Hedges'  $g = -0.3$ , 95% CI:  $-0.53 - -0.086$ ,  $I^2 = 27$ ,  $n = 11$ ,  $p < 0.05$ , 72.7% children; Figure 5). No significant effect size was observed for GABA in the cerebellum ( $n = 2$ ; 100% children), frontal ( $n = 7$ ; 50% children), temporal ( $n = 6$ ; 60% children), prefrontal ( $n = 4$ ; 25% children), parietal ( $n = 6$ ; 83% children) and occipital regions ( $n = 6$ ; 100% children).

To explore this in more detail we grouped GABA data using brain region grouping 2, where limbic regions are differentiated. We observed a negative effect size in the ACC only, however this did not withstand a more stringent p value threshold of  $p < 0.01$  due to  $df < 4$  (Supplementary figure 2; Hedges'  $g = -0.4$ , 95% CI:  $-0.74 - -0.065$ ,  $I^2 = 0$ ,  $n = 5$ ,  $p < 0.05$ ,  $df = 2.6$ ). Note for other limbic regions (thalamus, striatum) study number was small ( $n=2-3$ ) and effect sizes were not significant.

*By age group* A significant negative group effect size was observed for studies measuring GABA in **children** (Hedges'  $g = -0.78$ , 95% CI:  $-1.40 - -0.16$ ,  $I^2 = 88$ ,  $n = 16$ ,  $p < 0.05$ ; Figure 6) but not **adults** (Hedges'  $g = -0.26$ , 95% CI:  $-0.72 - 0.2$ ,  $I^2 = 75$ ,  $n = 11$ ,  $p = 0.24$ ; Figure 6).

*By age group and brain region* When studies measuring GABA were split by brain region (grouping 1) and age group (Figure 7 – 13), a significant negative group effect size was observed for **limbic** GABA in **children** (Hedges'  $g = -0.35$ , 95% CI:  $-0.69 - -0.011$ ,  $I^2 = 45$ ,  $n = 8$ ,  $p < 0.05$ ; Figure 9). This was also observed in **adults** when  $p < 0.05$ , however due to  $df < 4$ , a conservative  $p < 0.01$  was used and significance was lost (Hedges'  $g = -0.2$ , 95% CI:  $-0.36 - -0.045$ ,  $I^2 = 0$ ,  $n = 3$ ,  $p = 0.0324$  ( $df < 4$ ); Figure 9).

For studies measuring GABA in the cerebellum ( $n_{\text{children}} = 2$ ,  $n_{\text{adult}} = 0$ ; Figure 7), frontal ( $n_{\text{children}} = 4$ ,  $n_{\text{adult}} = 3$ ; Figure 8), prefrontal ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 3$ ; Figure 10), parietal ( $n_{\text{children}} = 5$ ,  $n_{\text{adult}} = 1$ ; Figure 11), temporal ( $n_{\text{children}} = 3$ ,  $n_{\text{adult}} = 0$ ; Figure 12) and occipital regions ( $n_{\text{children}} = 3$ ,  $n_{\text{adult}} = 3$ ; Figure 13) no significant effect size was observed for either age group.

To explore significant limbic findings in more detail, we grouped data by brain region grouping 2 (where limbic regions are differentiated) and age group. Again, the ACC was the only limbic region with sufficient study numbers to explore age effects. We observed a significant negative effect size for ACC GABA concentrations in children (Hedges'  $g = -0.5$ , 95% CI:  $-0.67 - 0.32$ ,  $I^2 = 0$ ,  $n = 4$ ,  $p < 0.01$ ; supplementary figure 3). In adults the study number was 1 (Hedges'  $g = -0.21$ , 95% CI:  $-0.64 - 0.22$ ,  $I^2 = 0$ ,  $n = 1$ ,  $p = 0.34$ )..

## **Choline**

*By brain region* A significant negative group effect size was found for studies measuring choline in the **parietal** regions (Hedges'  $g = -0.22$ , 95% CI:  $-0.44 - -0.006$ ,  $I^2 = 0$ ,  $n = 5$ ,  $p < 0.05$ , 25.7% children; Figure 5). No significant effect size was observed for choline in the

limbic (n=24; 44.1% children), cerebellum (n = 7; 75% children), frontal (n=1; 100% children), and prefrontal (n=9; 46.7 % children).

*By age group* We observed no significant effect size for studies measuring choline in adult only (Hedges'  $g = 0.08$ , 95% CI:  $-0.12 - 0.28$ ,  $I^2 = 13$ ,  $n = 15$ ,  $p = 0.40$ ; Figure 6) or children only subgroups (Hedges'  $g = 0.06$ , 95% CI:  $-0.14 - 0.27$ ,  $I^2 = 44$ ,  $n = 13$ ,  $p = 0.50$ ; Figure 6).

*By age group and brain region* When studies measuring choline were split by brain region and age group (Figure 7 - 13), no significant effect size was observed for any of the brain regions studied: cerebellum ( $n_{\text{children}} = 5$ ,  $n_{\text{adult}} = 2$ ; Figure 8), frontal ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 0$ ; Figure 8), limbic ( $n_{\text{children}} = 12$ ,  $n_{\text{adult}} = 13$ ; Figure 9), prefrontal ( $n_{\text{children}} = 3$ ,  $n_{\text{adult}} = 6$ ; Figure 10) or parietal ( $n_{\text{children}} = 2$ ,  $n_{\text{adult}} = 4$ ; Figure 11).

## **Creatine**

*By brain region* A significant negative group effect size was observed for studies measuring creatine in the cerebellum (Hedges'  $g = -0.23$ , 95% CI:  $-0.47 - -0.039$ ,  $I^2 = 0$ ,  $n = 4$ ,  $p < 0.05$ , 50% children; Figure 6). No significant effect size was observed for creatine in the limbic ( $n = 21$ ; 46.4 % children), frontal ( $n = 1$ ; 100% children), prefrontal ( $n = 10$ ; 41.7 % children) and parietal regions ( $n = 6$ ; 12.5 % children). For the temporal region  $n = 1$ .

*By age group* No significant effect size was observed for studies measuring creatine in adult only (Hedges'  $g = 0.03$ , 95% CI:  $-0.23 - 0.28$ ,  $I^2 = 68$ ,  $n = 15$ ,  $p = 0.82$ ; Figure 6) or children only subgroups (Hedges'  $g = -0.38$ , 95% CI:  $-0.84 - 0.08$ ,  $I^2 = 79$ ,  $n = 13$ ,  $p = 0.097$ ; Figure 6).

*By age group and brain region* When studies measuring creatine were split by brain region and age group (Figure 7 - 13), we observed a significant negative group effect size for **parietal** creatine in **adults** (Hedges'  $g = -0.19$ , 95% CI:  $-0.37 - -0.014$ ,  $I^2 = 0$ ,  $n = 4$ ,  $p < 0.05$ ; Figure 10) but not in **children** (Hedges'  $g = 0.00$ , 95% CI:  $-0.59 - 0.59$ ,  $I^2 = NA$ ,  $n = 2$ ,  $p = 0.25$ ; Figure 10).

No significant effect sizes were observed for the cerebellum ( $n_{\text{children}} = 2$ ,  $n_{\text{adult}} = 2$ ; Figure 7), frontal ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 0$ ; Figure 8), limbic ( $n_{\text{children}} = 11$ ,  $n_{\text{adult}} = 11$ ; Figure 9), prefrontal ( $n_{\text{children}} = 3$ ,  $n_{\text{adult}} = 7$ ; Figure 10) or temporal regions ( $n_{\text{children}} = 0$ ,  $n_{\text{adult}} = 1$ ; Figure 12) in children or adults.

### **Glutamate and Glx**

*By brain region* No significant effect size was observed for Glu in any region studied; limbic ( $n = 11$ ; 40.7% children), cerebellum ( $n = 3$ ; 66.7% children), frontal ( $n = 2$ ; 100% children), prefrontal ( $n = 2$ ; 0% children). For the occipital and temporal regions  $n = 1$ .

Similarly, for Glx, no single brain region had a significant group effect size; limbic ( $n = 19$ ; 40.7% children), cerebellum ( $n = 3$ , 66.7% children), frontal ( $n = 3$ , 33.3% children), temporal ( $n = 2$ , 0% children), prefrontal ( $n = 7$ , 14.3% children), parietal ( $n = 7$ , 30% children) and occipital ( $n = 4$ , 16.7% children).

*By age group* No significant effect size was observed for studies measuring Glu in adults (Hedges  $G' = -0.29$ , 95% CI:  $-0.78 - 0.21$ ,  $I^2 = 66$ ,  $n = 4$ ,  $p = 0.451$ ; Figure 6) or children (Hedges'  $g = 0.30$ , 95% CI:  $-0.031 - 0.63$ ,  $I^2 = 69$ ,  $n = 9$ ,  $p = 0.06$ ; Figure 6). The same was observed for Glx (Hedges'  $g_{\text{children}} = -0.27$ , 95% CI:  $-0.83 - 0.28$ ,  $I^2 = 84$ ,  $n = 12$ ,  $p = 0.30$  and Hedges'  $g_{\text{adults}} = -0.04$ , 95% CI:  $-0.18 - 0.25$ ,  $I^2 = 62$ ,  $n = 17$ ,  $p = 0.74$ ).

*By age group and brain region* For Glu no significant effect size was observed for children or adults in any brain region studied: cerebellum ( $n_{\text{children}} = 2$ ,  $n_{\text{adult}} = 1$ ; Figure 7), frontal ( $n_{\text{children}} = 2$ ,  $n_{\text{adult}} = 0$ ; Figure 8), limbic ( $n_{\text{children}} = 8$ ,  $n_{\text{adult}} = 3$ ; Figure 9), prefrontal ( $n_{\text{children}} = 0$ ,  $n_{\text{adult}} = 2$ ; Figure 10), temporal ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 0$ ; Figure 12) and occipital ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 0$ ; Figure 13).

The same was observed for Glx: cerebellum ( $n_{\text{children}} = 2$ ,  $n_{\text{adult}} = 1$ ; Figure 7), frontal ( $n_{\text{children}} = 0$ ,  $n_{\text{adult}} = 1$ ; Figure 8), limbic ( $n_{\text{children}} = 10$ ,  $n_{\text{adult}} = 10$ ; Figure 9), prefrontal ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 7$ ; Figure 10), parietal ( $n_{\text{children}} = 3$ ,  $n_{\text{adult}} = 4$ ; Figure 11), temporal ( $n_{\text{children}} = 0$ ,  $n_{\text{adult}} = 2$ ; Figure 12) and occipital ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 3$ ; Figure 13).

## **GSH**

*By brain region* When studies measuring GSH were split by brain region (Figure 5), no single brain region studied had a significant group effect size, this included limbic ( $n = 2$ , 0% children), frontal ( $n = 1$ , 0% children), prefrontal ( $n = 2$ , 0% children) and occipital regions ( $n = 1$ , 0% children).

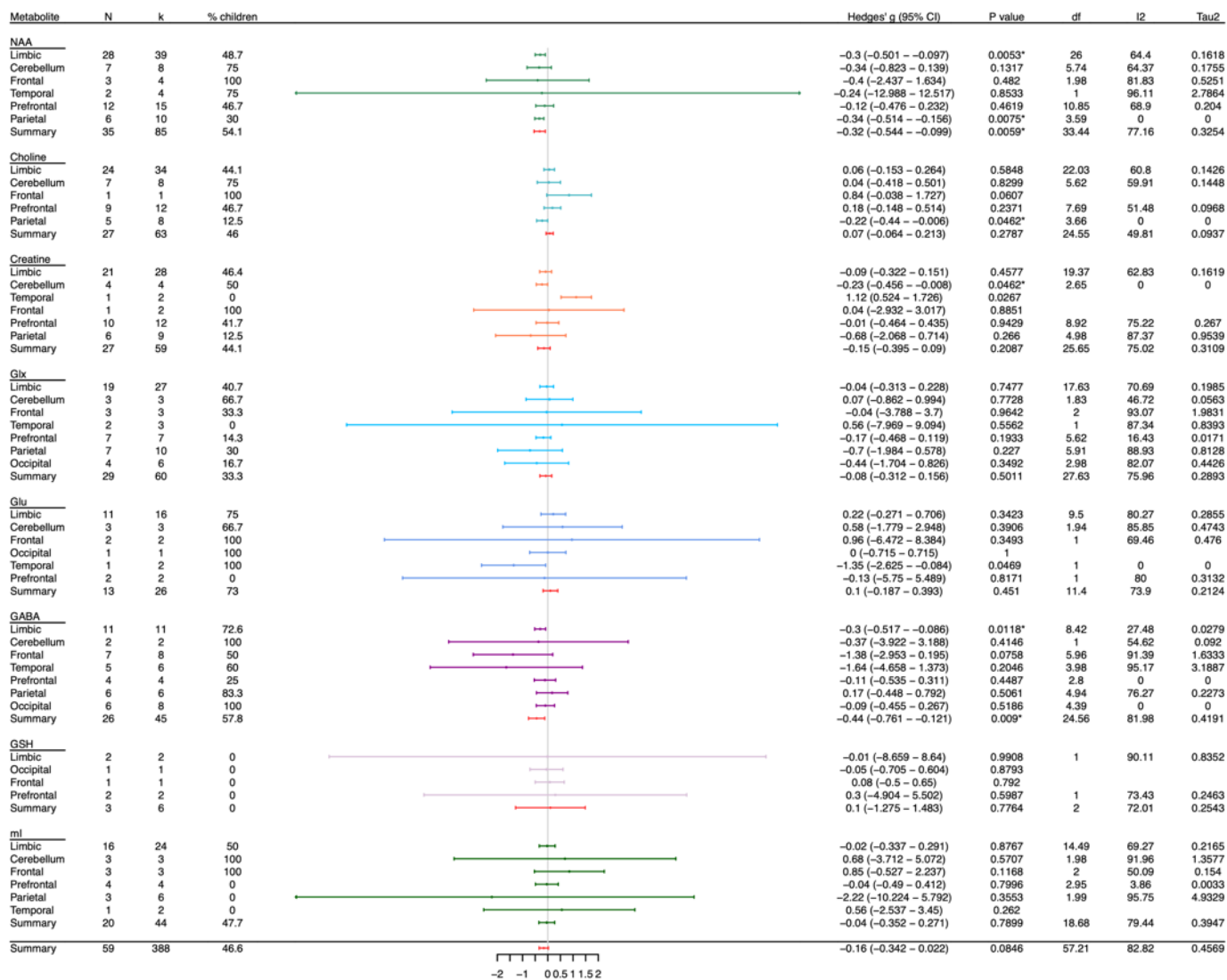
*By age group* GSH has only been studied in adult cohorts, for which no significant effect size was observed (Hedges'  $g = 0.1$ , 95% CI: -1.28 - 1.48,  $I^2 = 72$ ,  $n = 3$ ,  $p = 0.78$ ).

## **ml**

*By brain region* For studies measuring ml, no brain region studied had a significant group effect size (Figure 5): limbic ( $n = 16$ , 40% children), cerebellum ( $n = 3$ , 100% children), frontal ( $n = 3$ , 100% children), temporal ( $n = 1$ , 0% children), prefrontal ( $n = 4$ , 0% children), parietal regions ( $n = 2$ , 0% children).

*By age group* When studies measuring ml were split by age group (Figure 6), no significant effect size was observed for adults (Hedges  $G' = -0.01$ , 95% CI: -0.19 - 0.20,  $I^2 = 40$ ,  $n = 10$ ,  $p = 0.93$ ) or children (Hedges'  $g = 0.09$ , 95% CI: -0.86 - 0.68,  $I^2 = 88$ ,  $n = 10$ ,  $p = 0.80$ ).

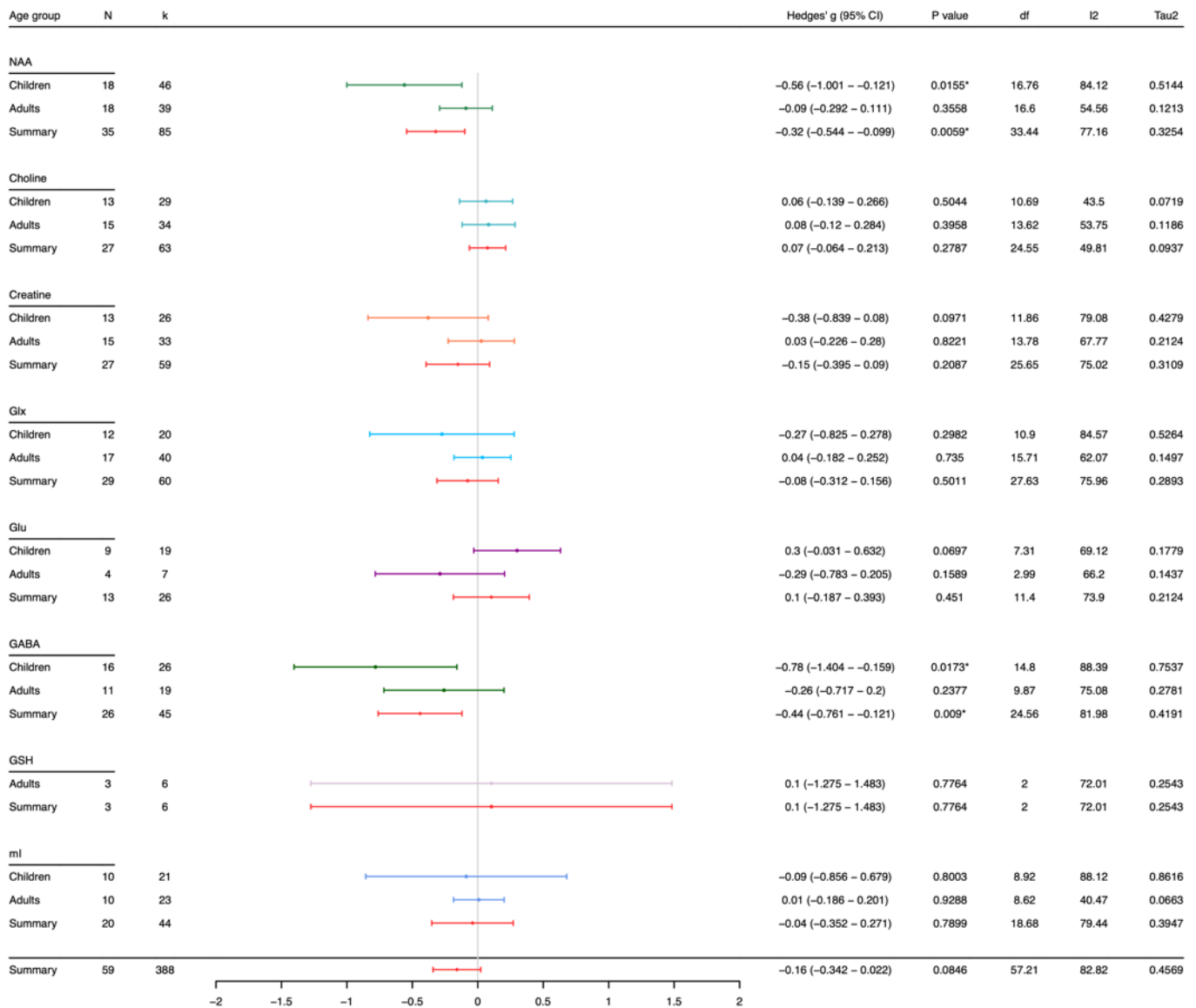
*By age group and brain region* When studies measuring ml were split by brain region and age group (Figure 7 – 13), we observed no significant effect size for any brain region in children or adults: cerebellum ( $n_{\text{children}} = 0$ ,  $n_{\text{adult}} = 3$ ; Figure 7), frontal ( $n_{\text{children}} = 2$ ,  $n_{\text{adult}} = 0$ ; Figure 8), limbic ( $n_{\text{children}} = 9$ ,  $n_{\text{adult}} = 7$ ; Figure 9), prefrontal ( $n_{\text{children}} = 0$ ,  $n_{\text{adult}} = 4$ ; Figure 10), parietal ( $n_{\text{children}} = 0$ ,  $n_{\text{adult}} = 2$ ; Figure 11) and temporal ( $n_{\text{children}} = 1$ ,  $n_{\text{adult}} = 1$ ; Figure 12).



**Figure 5.** Summary forest plot for data grouped by metabolite and brain region.

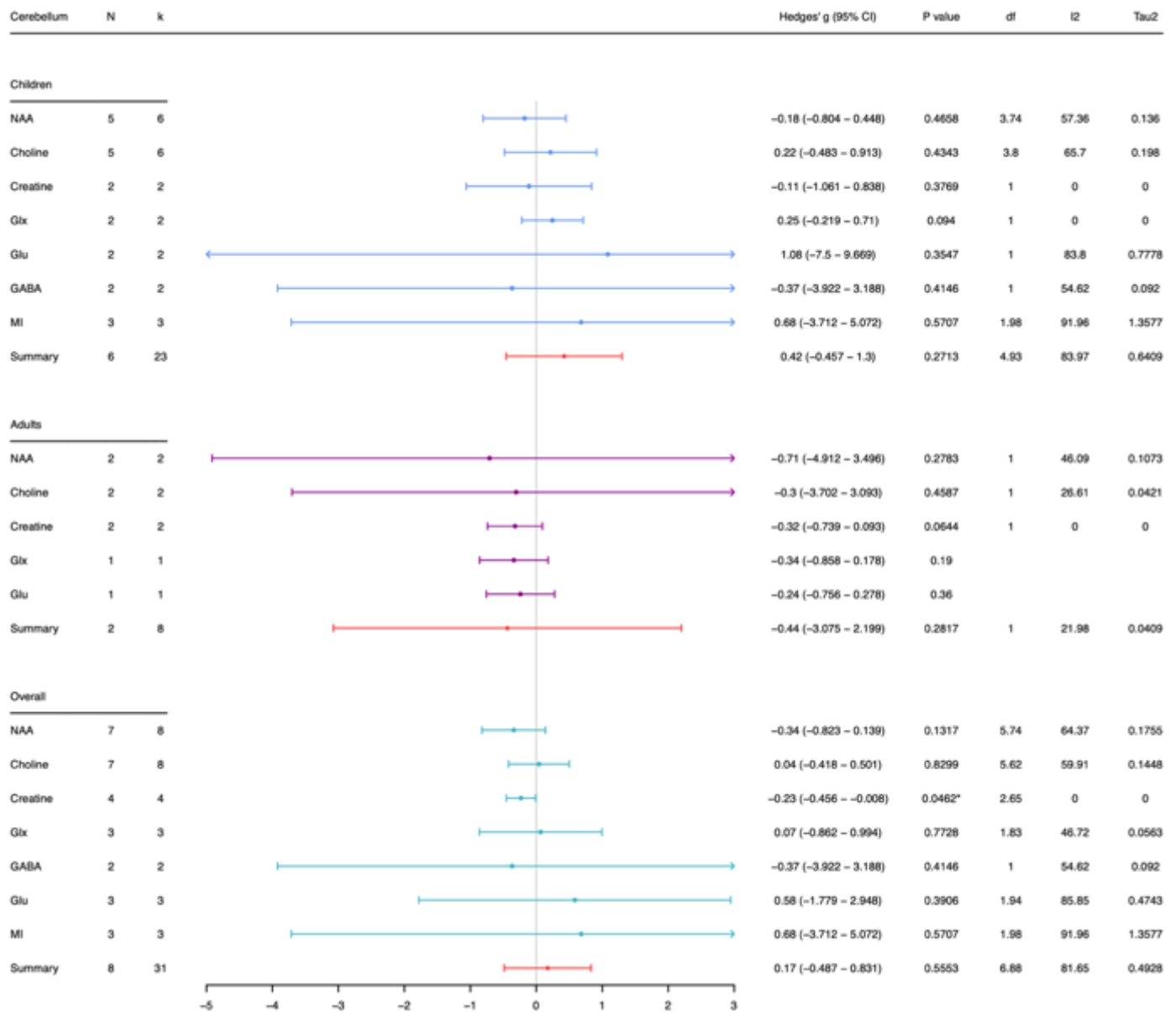
N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



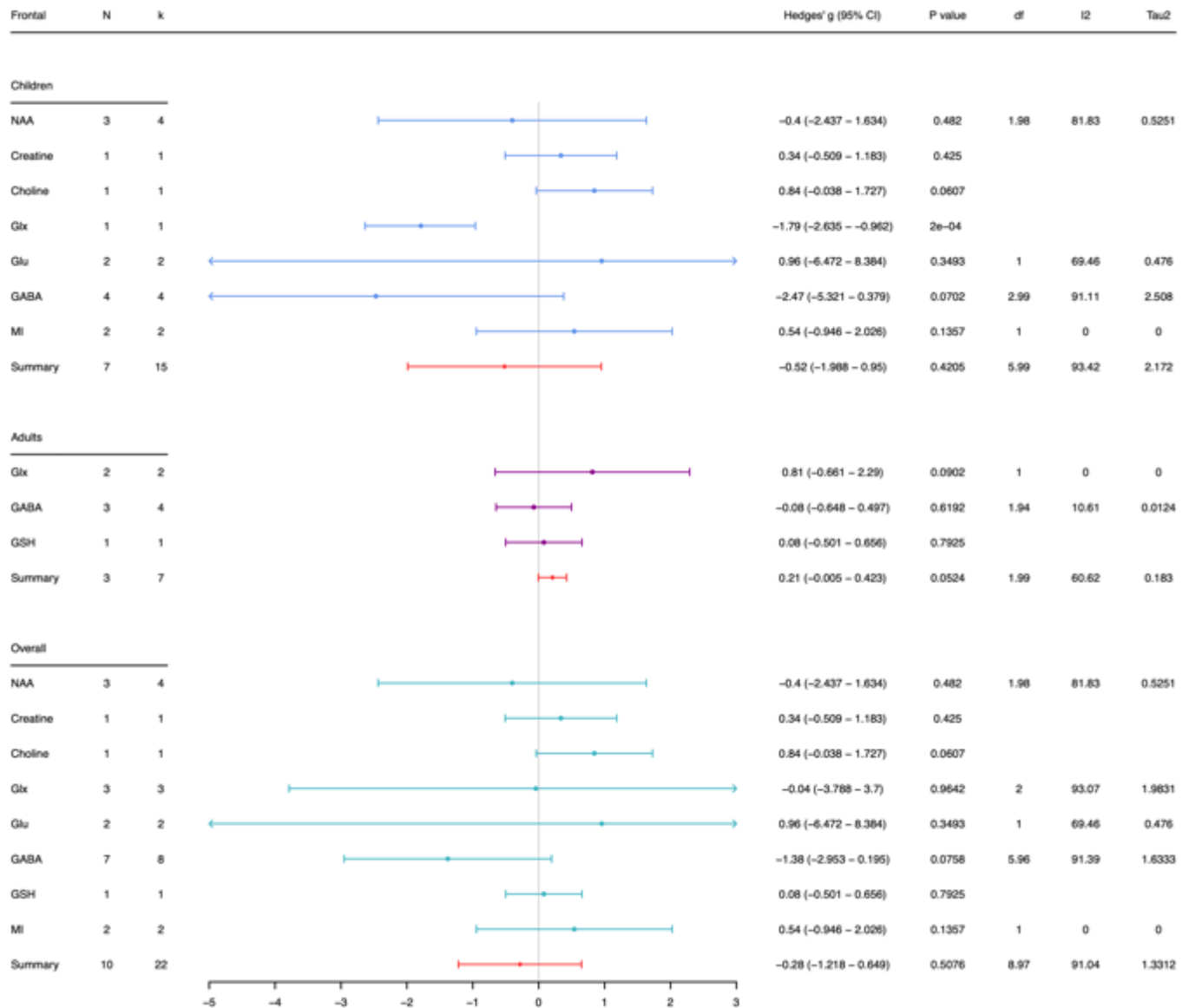


**Figure 6.** Summary forest plot for data by metabolite and age group.

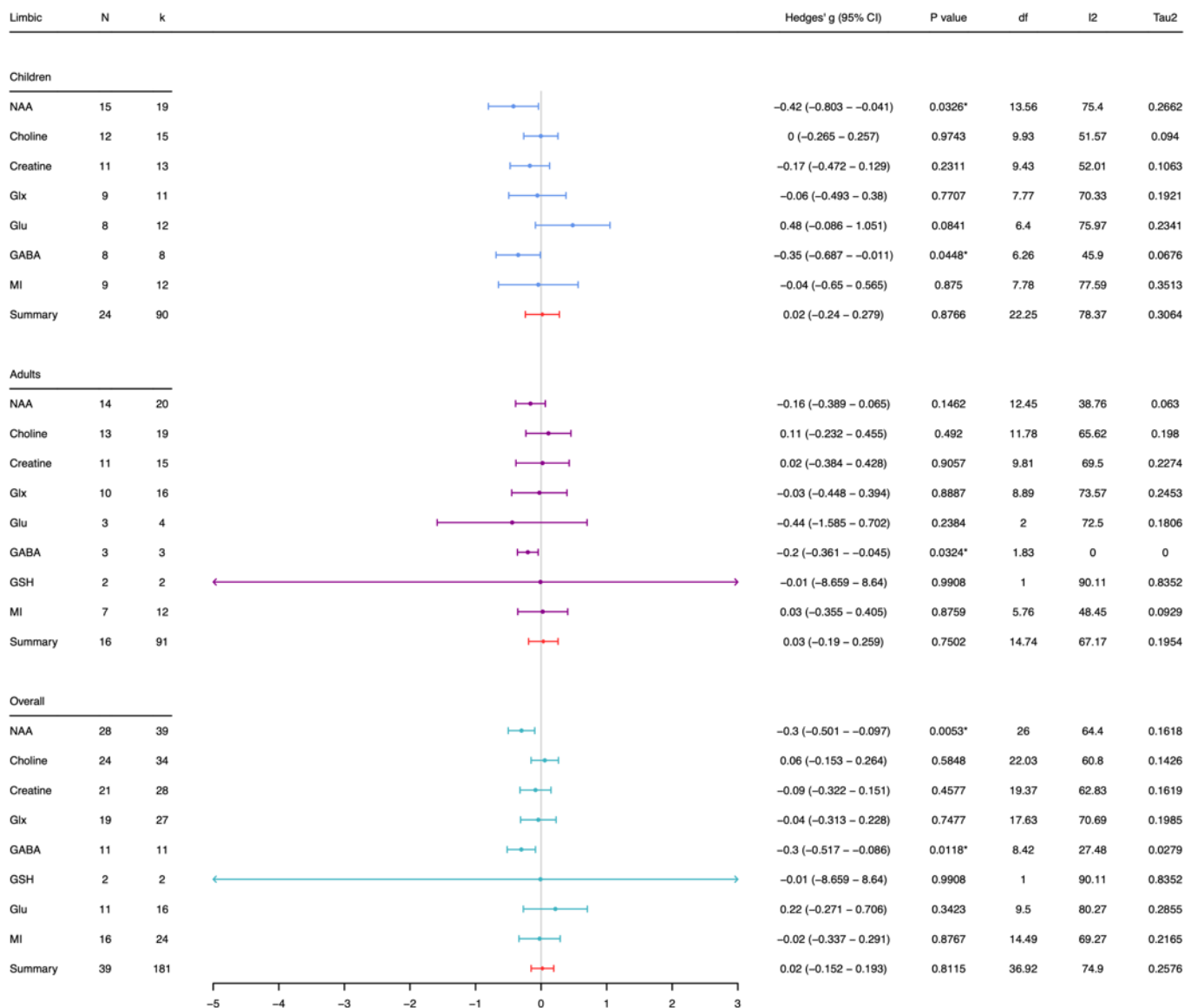
N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



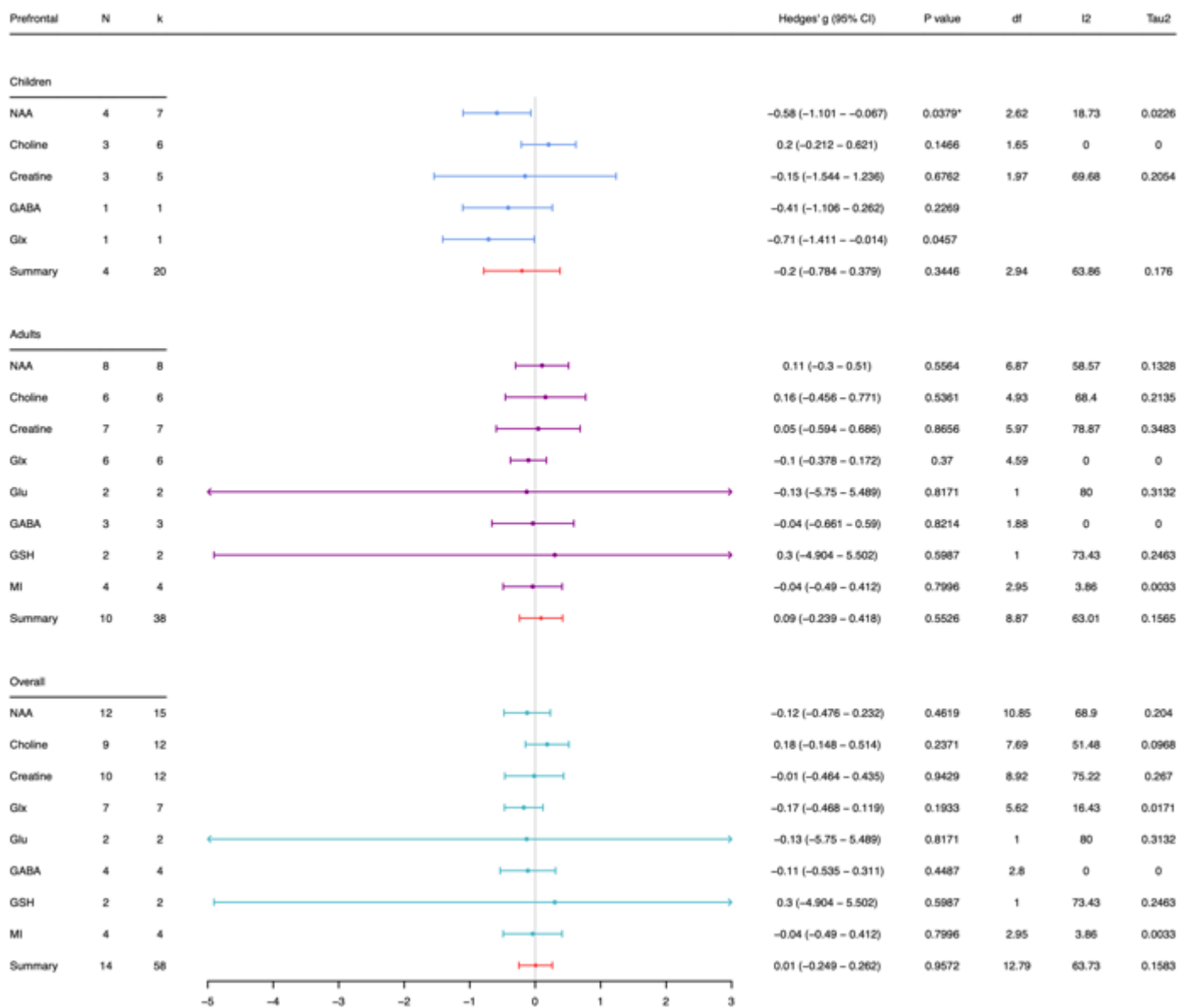
**Figure 7.** Summary forest plot findings in the cerebellum split by metabolite and age group (children, adults, and pooled ages (overall)). N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



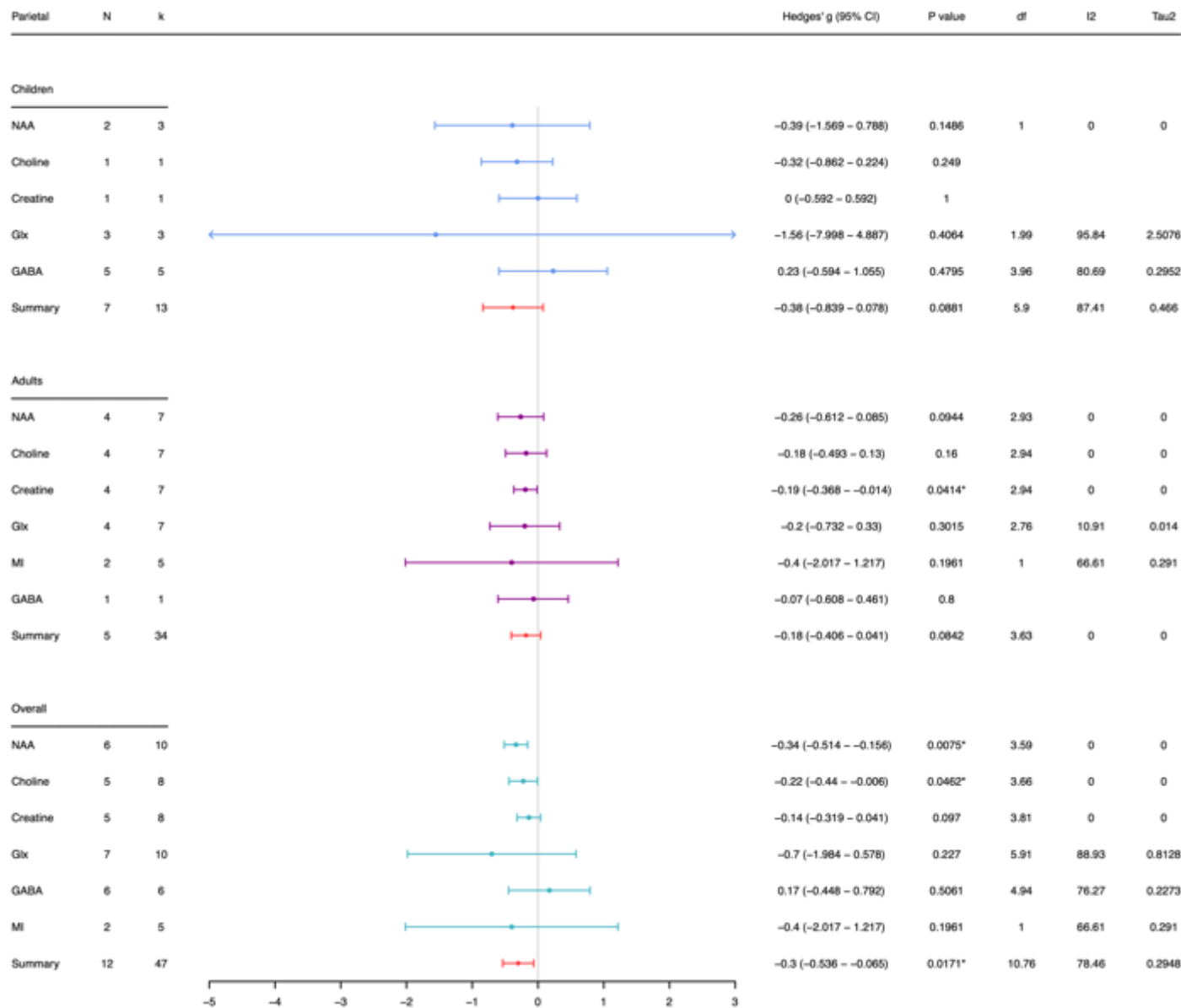
**Figure 8.** Summary forest plot findings in the frontal region split by metabolite and age group (children, adults, and pooled ages (overall)). N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



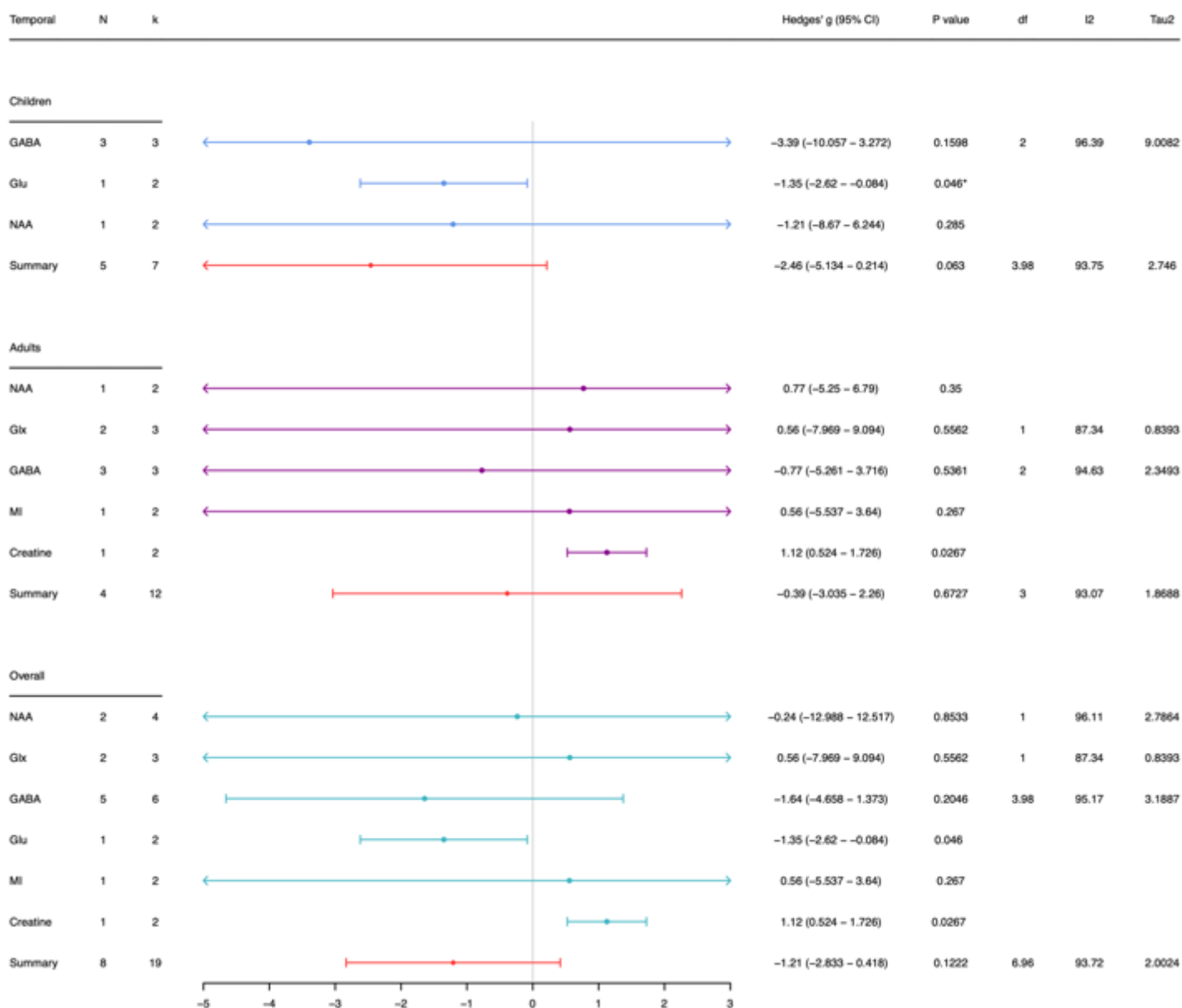
**Figure 9.** Summary forest plot findings in the limbic regions split by metabolite and age group (children, adults, and pooled ages (overall)). N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



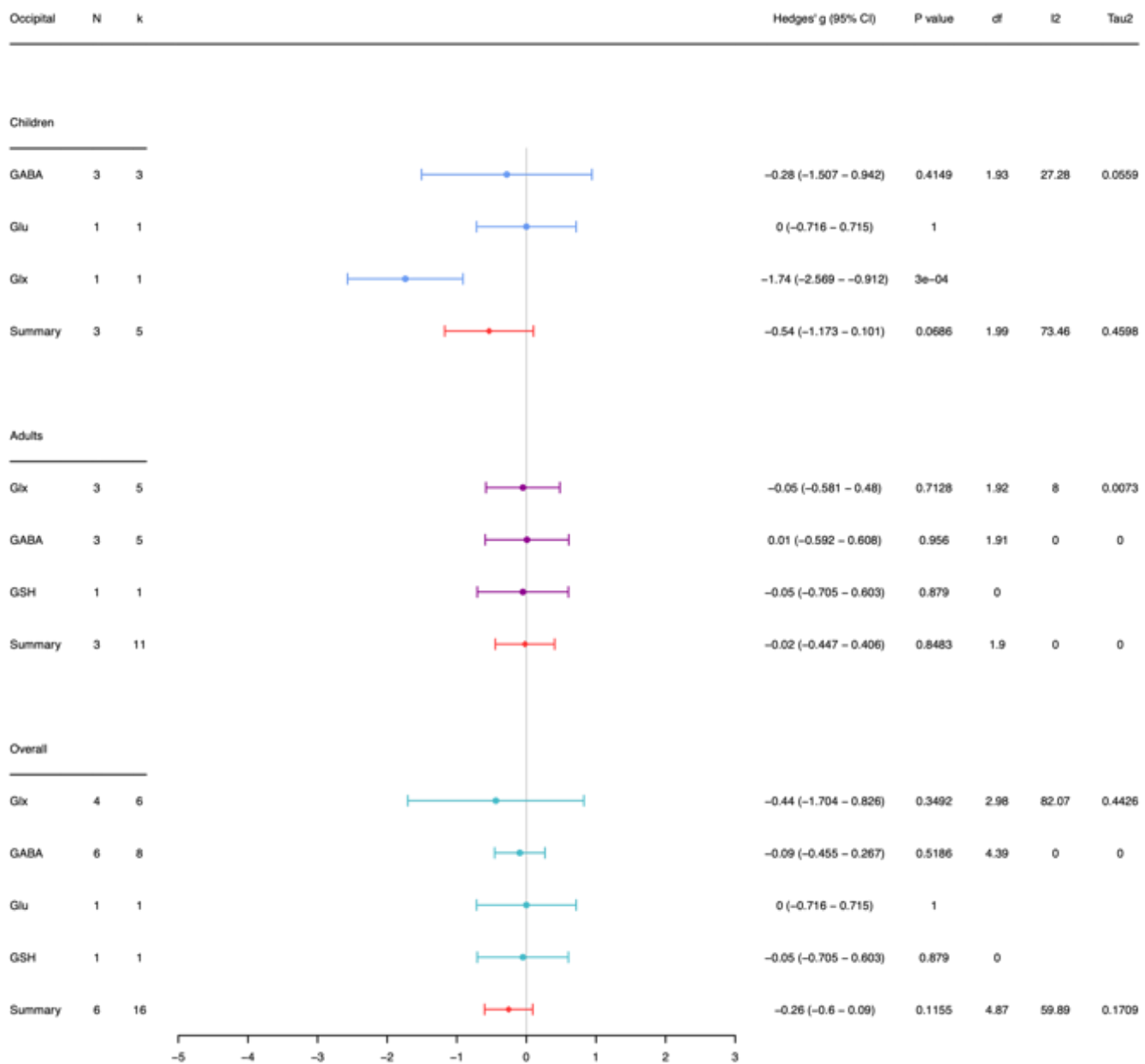
**Figure 10.** Summary forest plot findings in the prefrontal region split by metabolite and age group (children, adults, and pooled ages (overall)). N: number of studies, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



**Figure 11.** Summary forest plot findings in the parietal region split by metabolite and age group (children, adults, and pooled ages (overall)). N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



**Figure 12.** Summary forest plot findings in the temporal region split by metabolite and age group (children, adults, and pooled ages (overall)). N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



**Figure 13.** Summary forest plot findings in the occipital region split by metabolite and age group (children, adults, and pooled ages (overall)). N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study heterogeneity). Hedge's g is reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the degrees of freedom < 4 for RVE t-tests.



### 3.6 Effect of demographic factors

The effect of MRS study demographics on effect size is presented below. When split by metabolite and brain region the number of studies per group was too small for meaningful conclusions, as such in most cases we observed the effect of these factors across all brain regions and ages and in most cases, metabolites to maximise power.

#### Sex

Data was grouped by cohort sex ratio. Overall, 43 case-control comparisons included females (of these 10 included at least a 1:3 female:male ratio, 7 included at least a 1:1 female:male ratio), 13 studies excluded females, and 5 failed to report participant sex. Note this total is greater than the number of studies, because in some cases studies included variable cohorts for different brain regions studied.

Across all brain regions, ages and metabolites no obvious differences in group effect size were observed between studies that included at least 1:1 female:male ratio in cohorts (Hedges'  $g_{50\%included} = -0.1$ , 95% CI: -0.35 - 0.15,  $I^2 = 39$ ,  $n = 7$ ,  $p = 0.36$ ) and those that did not (Hedges'  $g_{50\%not} = -0.11$ , 95% CI: -0.30 - 0.078,  $I^2 = 49$ ,  $n = 51$ ,  $p = 0.24$ ). For studies that included 1:3 female:male ratio in cohorts, an effect size close to 0 was observed (Hedges'  $g_{1:3included} = -0.02$ , 95% CI: -0.19 - 0.24,  $I^2 = 71$ ,  $n = 21$ ,  $p = 0.82$ ; Figure 14). For studies that did not include at least a 1:3 ratio of female:males, a greater, negative effect size was observed. (Hedges'  $g_{1:3not} = -0.21$ , 95% CI: -0.47 - 0.042,  $I^2 = 84$ ,  $n = 34$ ,  $p = 0.098$ ; Figure 14).

When data were grouped by metabolite, for studies measuring GABA (supplementary figure 4), those that included at least a 1:1 ratio of females to males in cohorts had a non-significant effect size close to zero (Hedges'  $g = -0.03$ , 95% CI:  $-0.33 - 0.26$ ,  $I^2 = 0$ ,  $n = 5$ ,  $p = 0.76$ ). Studies measuring GABA that did not include at least a 1:1 ratio of females to males in cohorts had a significant and negative effect size (Hedges'  $g = -0.54$ , 95% CI:  $-1.00 - -0.082$ ,  $I^2 = 86$ ,  $n = 19$ ,  $p < 0.05$ ). Note study numbers per group were too small to assess the effect of sex on limbic GABA in particular.

Studies measuring GABA that included at least a 1:3 female to male ratio in cohorts (supplementary figure 4) had a non-significant negative size (Hedges'  $g = -0.25$ , 95% CI:  $-0.81 - 0.32$ ,  $I^2 = 80$ ,  $n = 9$ ,  $p = 0.35$ ). Studies measuring GABA that did not include at least a 1:3 female to male ratio in cohorts had a significant and negative effect size (Hedges'  $g = -0.59$ , 95% CI:  $-1.12 - -0.051$ ,  $I^2 = 85$ ,  $n = 14$ ,  $p < 0.05$ ).

For studies measuring NAA (supplementary figure 5), those that included at least a 1:1 ratio of females to males in cohorts had a non-significant negative effect size (Hedges'  $g = -0.55$ , 95% CI:  $-5.84 - 4.75$ ,  $I^2 = 86$ ,  $n = 2$ ,  $p = 0.42$ ). Studies measuring NAA that did not include at least a 1:1 ratio of females to males in cohorts had a significant and negative effect size (Hedges'  $g = -0.24$ , 95% CI:  $-0.42 - -0.054$ ,  $I^2 = 68$ ,  $n = 30$ ,  $p < 0.05$ ).

Studies assessing NAA that included a minimum ratio of 1:3 females to males in cohorts (supplementary figure 5) had a non-significant negative effect size (Hedges'  $g = -0.24$ , 95% CI:  $-0.65 - 0.16$ ,  $I^2 = 75$ ,  $n = 11$ ,  $p = 0.21$ ). Studies measuring NAA that did not include at least

a ratio of 1:3 females to males in cohorts had a significant and negative effect size (Hedges'  $g = -0.24$ , 95% CI:  $-0.45 - -0.029$ ,  $I^2 = 67$ ,  $n = 19$ ,  $p < 0.05$ ).

## IQ

Data was grouped by cohort IQ. Overall, 44 case-control comparisons excluded individuals with IQ < 70, 6 studies did not exclude individuals with IQ < 70, 9 studies failed to report IQ of participants.

Across all metabolites, brain regions and age groups, studies that included individuals with an IQ < 70 had a non-significant negative effect size (Hedges'  $g = -0.47$ , 95% CI:  $-1.39 - 0.44$ ,  $I^2 = 87$ ,  $n = 6$ ,  $p = 0.24$ ; Figure 14). Those that excluded individuals with IQ < 70 had a non-significant effect size close to 0 (Hedges'  $g = -0.09$ , 95% CI:  $-0.31 - 0.14$ ,  $I^2 = 83$ ,  $n = 44$ ,  $p = 0.43$  respectively; Figure 14).

When data were grouped by metabolite, for GABA (supplementary figure 4), studies excluding individuals with an IQ < 70 had a significant and negative effect size (Hedges'  $g = -0.27$ , 95% CI:  $-0.56 - -0.022$ ,  $I^2 = 74$ ,  $n = 19$ ,  $p < 0.05$ ). Studies including individuals with an IQ < 70 had a non-significant negative effect size (Hedges'  $g = -1.16$ , 95% CI:  $-4.36 - 2.05$ ,  $I^2 = 91$ ,  $n = 3$ ,  $p = 0.92$ ). Note the study number discrepancy between these groups.

For studies measuring NAA (supplementary figure 5), those excluding individuals with an IQ < 70 had a non-significant and negative effect size (Hedges'  $g = -0.23$ , 95% CI:  $-0.59 - 0.13$ ,  $I^2 = 80$ ,  $n = 22$ ,  $p = 0.19$ ). The same was observed for studies including individuals with an IQ < 70 (Hedges'  $g = -0.37$ , 95% CI:  $-0.856 - 0.12$ ,  $I^2 = 73$ ,  $n = 5$ ,  $p = 0.10$ ).

## Psychotropic medication

Data were grouped by reported exclusion or inclusion based on psychoactive medication. 16 studies reported the exclusion of participants taking psychotropic medication at the time of scan. 28 did not exclude individuals on psychotropic medication at the time of scan. These studies reported varying proportions of participants on varying psychotropic medications, the most common being stimulants, anti-depressants, and anti-psychotics. 15 studies failed to report the medication status of participants.

Across all brain regions, ages and metabolites, studies excluding psychotropic medication use had a non-significant positive group effect size (Hedges'  $g = 0.14$ , 95% CI:  $-0.077 - 0.35$ ,  $I^2 = 72$ ,  $n = 16$ ,  $p = 0.18$ ). Studies including individuals on psychotropic medications at time of scan had a non-significant negative group effect size (Hedges'  $g = -0.31$ , 95% CI:  $-0.66 - 0.028$ ,  $I^2 = 85$ ,  $n = 28$ ,  $p = 0.070$ ; Figure 14).

When data were grouped by metabolite, for GABA (supplementary figure 4), studies excluding psychotropic medication use had a non-significant effect size at zero (Hedges'  $g = 0.00$ , 95% CI:  $-0.40 - 0.42$ ,  $I^2 = 0$ ,  $n = 3$ ,  $p = 0.92$ ). Studies that included individuals on psychotropic medications at the time of scan had a significant negative group effect size (Hedges'  $g = -0.45$ , 95% CI:  $-0.84 - 0.050$ ,  $I^2 = 80$ ,  $n = 18$ ,  $p < 0.05$ ).

For NAA (supplementary figure 5), studies excluding psychotropic medication use had a non-significant negative effect size (Hedges'  $g = -0.17$ , 95% CI:  $-0.46 - 0.12$ ,  $I^2 = 68$ ,  $n = 12$ ,  $p =$

0.23). The same was observed for studies including individuals taking psychotropic medications (Hedges'  $g = -0.45$ , 95% CI:  $-1.34 - 0.44$ ,  $I^2 = 87$ ,  $n = 10$ ,  $p = 0.28$ ).

### **Sedation**

Data was grouped by the reported use of sedation. In 6 studies, all or some of the participants received sedation during the MRI procedures: (Bejjani et al., 2012: 1 participant, intravenous propofol; Goji et al., 2017: all participants, 0.5mL/kg body weight of triclofos sodium; Harada et al., 2011: half of participants, 0.2 mL/kg body weight of triclofos sodium; Hashimoto et al., 1997: all children under 6,  $n = ?$ , triclofos sodium; Ito et al., 2017, all participants, 0.5mL/kg body weight of triclofos sodium; Mori et al., 2013: "most" participants, 0.5mL/kg body weight of triclofos sodium). 53 studies did not report use of sedation.

No obvious differences in group effect size was observed between studies that reported use of sedation (Hedges'  $g = -0.19$ , 95% CI:  $-0.45 - 0.06$ ,  $I^2 = 60$ ,  $n = 6$ ,  $p = 0.11$ ) and those that did not (Hedges'  $g = -0.16$ , 95% CI:  $-0.37 - 0.051$ ,  $I^2 = 84$ ,  $n = 53$ ,  $p = 0.13$ ; Figure 14; across all brain regions, ages and metabolites), although study number disparity between groups may contribute to this.

When data were grouped by metabolite, for GABA (supplementary figure 4), studies that reported use of sedation had a significant and negative effect size (Hedges'  $g = -0.52$ , 95% CI:  $-0.80 - -0.25$ ,  $I^2 = 39$ ,  $n = 4$ ,  $p < 0.05$ ). The same was observed for studies that did not report use of sedation (Hedges'  $g = -0.44$ , 95% CI:  $-0.83 - -0.041$ ,  $I^2 = 83$ ,  $n = 22$ ,  $p < 0.05$ ).

For studies measuring NAA (supplementary figure 5), studies that reported use of sedation had a negative effect size (Hedges'  $g = -0.15$ , 95% CI:  $-0.59 - 0.29$ ,  $I^2 = 63$ ,  $n = 7$ ,  $p = 0.43$ ).

Studies that did not report use of sedation had a significant and negative effect size (Hedges'  $g = -0.38$ , 95% CI:  $-0.66 - -0.098$ ,  $I^2 = 79$ ,  $n = 28$ ,  $p < 0.05$ ).

### **Reported diagnosis**

Studies report various methods for autism classification of the included autism cohort, including classification according to the American Psychiatric Association's Diagnostic and Statistical Manual, fourth addition (DSM-4; American Psychiatric Association, 2000), fifth addition (DSM-5; *Diagnostic and Statistical Manual of Mental Disorders*, 2013) and the International Classification of Diseases 10<sup>th</sup> Revision (ICD-10, World Health Organization, 1993). Studies used various approaches for diagnosis including the Autism Diagnostic Interview—Revised ADI-R (Lord et al., 1994), the Autism Diagnostic Observation Schedule 1 (ADOS1; Gotham et al., 2007) and 2 (Lord et al., 2012); see Table 4 for specifics per study.

DSM-5 (American Psychiatric Association., 2013) criteria abandoned the sub-grouping of its predecessor DSM-4 (American Psychiatric Association., 1994), including autistic disorder (AD) and Aspergers syndrome (AS), under 'ASD' and facilitating the dual diagnosis of autism with other neurodevelopmental disorders such as ADHD (Colleen & Stone, 2014). As many MRS studies predate this change in classification, studies observing AD (as classified by DSM-4 and/or ICD-10), AS (as classified by DSM-4 and/or ICD-10), and ASD (as classified by DSM-5/ADI-R) were included in this meta-analysis.

Data were grouped by reported diagnosis/classification of autism cohorts (Table 4). 47 studies reported on a cohort classified as having ASD, inferred via the inclusion of multiple DSM-4/ICD-10 subgroups (AD & AS), and/or explicitly stating this classification in the study methods (DSM-5/ADI-R). Of note, 8 studies used DSM-4 classification but included individuals with AS or AD and so were classified as 'ASD' (see Table 4). 8 studies reported on individuals classified as having AD; this was inferred by the exclusion of any other developmental disorders (including AS) under the DSM-4 **and** the explicit statement of the inclusion of this classification only (DSM-4/ICD-10). 4 studies reported data from individuals with an AS classification, this being explicitly stated (DSM-4 and/or ICD-10).

Across all metabolites, brain regions and ages, studies reporting on ASD had a negative group effect size (Hedges'  $g = -0.19$ , 95% CI:  $-0.40 - 0.02$ ,  $I^2 = 83$ ,  $n = 47$ ,  $p = 0.076$ ; Figure 14). For studies reporting on AS, a group effect size close to 0 was observed (Hedges'  $g = 0.08$ , 95% CI:  $-0.88 - 0.12$ ,  $I^2 = 57$ ,  $n = 4$ ,  $p = 0.69$ ). The same was observed for studies reporting on AD (Hedges'  $g = -0.09$ , 95% CI:  $-0.78 - 0.60$ ,  $I^2 = 87$ ,  $n = 8$ ,  $p = 0.76$ ; Figure 14).

When data were grouped by metabolite, for GABA (supplementary figure 4), studies that reported on ASD had a significant and negative group effect size (Hedges'  $g = -0.41$ , 95% CI:  $-0.76 - -0.048$ ,  $I^2 = 80$ ,  $n = 24$ ,  $p < 0.05$ ). Studies reporting on AS had a non-significant, negative group effect size (Hedges'  $g = -0.33$ , 95% CI:  $-4.0 - 3.30$ ,  $I^2 = 37$ ,  $n = 2$ ,  $p = 0.46$ ). The same was observed for studies reporting on AD (Hedges'  $g = -2.93$ , 95% CI:  $-26.62 - 20.77$ ,  $I^2 = 94$ ,  $n = 2$ ,  $p = 0.36$ ).

For NAA (supplementary figure 5), studies that reported on ASD had a significant and negative effect size (Hedges'  $g = -0.31$ , 95% CI:  $-0.61 - -0.015$ ,  $I^2 = 78$ ,  $n = 24$ ,  $p < 0.05$ ). The same was observed for studies reporting AD (Hedges'  $g = -0.61$ , 95% CI:  $-0.97 - -0.25$ ,  $I^2 = 70$ ,  $n = 7$ ,  $p < 0.05$ ). For studies reporting AS, we observe a non-significant positive group effect size (Hedges'  $g = 0.16$ , 95% CI:  $-0.49 - 0.82$ ,  $I^2 = 67$ ,  $n = 4$ ,  $p = 0.48$ ).

### 3.7 Effect of MRS methods

The effect of MRS parameters on effect size was also assessed.

#### Quantification methods & tissue correction

Data was grouped by metabolite quantification method: 48 studies reported metabolite concentrations in institutional units (i.u.; water scaled). 31 of these performed some form of tissue correction. 14 studies reported metabolite concentrations as creatine ratios only. In 3 studies, edited metabolite concentrations (GABA+/GSH) were recorded as creatine ratios, while unedited metabolites (NAA etc) were reported in institutional units (Cochran et al., 2015; Fung et al., 2021; Sapey-Triomphe et al., 2021).

Across all brain regions, ages and metabolites, the group effect size of studies reporting metabolite concentrations in institutional units was 0 (Hedges'  $g = -0.06$ , 95% CI:  $-0.21 - 0.088$ ,  $I^2 = 77$ ,  $n = 48$ ,  $p = 0.42$ ; Figure 14). The group effect size of studies reporting metabolite concentrations as creatine ratios was negative (Hedges'  $g = -0.51$ , 95% CI:  $-1.15 - 0.12$ ,  $I^2 = 89$ ,  $n = 15$ ,  $p = 0.10$ ; Figure 14).



When data were grouped by metabolite of interest (supplementary figure 5), studies reporting water-scaled NAA concentrations in i.u. had a significant negative group effect size (Hedges'  $g = -0.29$ , 95% CI:  $-0.56 - -0.031$ ,  $I^2 = 77$ ,  $n = 27$ ,  $p < 0.05$ ). Studies reporting NAA concentrations as creatine ratios had a non-significant negative group effect size (Hedges'  $G = -0.42$ , 95% CI:  $-0.57 - 0.12$ ,  $I^2 = 79$ ,  $n = 8$ ,  $p = 0.11$ ).

Studies reporting GABA concentration as a creatine ratio had a significant negative group effect size (Hedges'  $g = -1.62$ , 95% CI:  $-2.97 - -0.27$ ,  $I^2 = 92$ ,  $n = 7$ ,  $p < 0.05$ ; supplementary figure 4). Studies reporting water-scaled GABA concentration in i.u. had a non-significant negative group effect size (Hedges'  $g = -0.11$ , 95% CI:  $-2.49 - 0.02$ ,  $I^2 = 37$ ,  $n = 19$ ,  $p = 0.09$ ).

The group effect size of studies that reported the use of tissue correction (Hedges'  $g = -0.05$ , 95% CI:  $-0.18 - 0.084$ ,  $I^2 = 72$ ,  $n = 29$ ,  $p = 0.45$ ; Figure 14) was very similar to that of those that did not report the use of tissue correction for water-scaled data (Hedges'  $g = -0.04$ , 95% CI:  $-0.51 - 0.44$ ,  $I^2 = 83$ ,  $n = 15$ ,  $p = 0.87$ ; Figure 14).

### **Scanner field strength**

Data was grouped by scanner field strength; 20 MRS studies were performed at 1.5 Tesla (T), 48 studies at 3T, and 1 study at 4T. No clear difference in group effect size was observed for the different scanner strengths (1.5, 3 & 4T; Figure 14; across all brain regions, metabolites, and ages).

When data were grouped by metabolite, for GABA (supplementary figure 4), one study at 1.5T had a non-significant negative effect size (Hedges'  $g = -2.93$ , 95% CI:  $-26.62 - 20.77$ ,  $I^2 =$

94,  $n = 1$ ,  $p = 0.36$ ). Most included studies measured GABA at 3T, for which a significant negative group effect size was observed (Hedges'  $g = -0.40$ , 95% CI:  $-0.73 - -0.058$ ,  $I^2 = 80$ ,  $n = 25$ ,  $p < 0.05$ ).

For studies that measured NAA at 1.5T (supplementary figure 5), we observe a significant negative effect size (Hedges'  $g = -0.47$ , 95% CI:  $-0.91 - -0.037$ ,  $I^2 = 84$ ,  $n = 19$ ,  $p < 0.05$ ). For studies measuring NAA at 3T, we observe a non-significant negative group effect size (Hedges'  $g = -0.16$ , 95% CI:  $-0.37 - 0.054$ ,  $I^2 = 52$ ,  $n = 15$ ,  $p = 0.13$ ).

### **MRS sequence**

Data was grouped by MRS acquisition pulse sequence used. 48 studies used non-editing MRS sequences; 37 studies used PRESS (Bottomley, 1987), 9 studies used Stimulated Echo Acquisition Mode (STEAM), 1 study used a two-dimensional J-resolved protocol (2D-JPRESS; (Joshi et al., 2013) and 1 study used semi-LASER (Pierce et al., 2021). 25 studies used editing MRS acquisition sequences; 21 studies used MEGA-PRESS; (Mescher et al., 1996, 1998), 2 studies used Hadamard Encoding and Reconstruction of MEGA-Edited Spectroscopy (HERMES; He et al., 2021; Saleh et al., 2016; Sapey-Triomphe et al., 2021) and 1 used MEGA-SPECIAL (Fung et al., 2021; M. Gu et al., 2018). Note 10 studies used both editing and non-editing sequences (Brix et al., 2015; Cochran et al., 2015; Drenthen et al., 2016; Harada et al., 2011; He et al., 2021; Ito et al., 2017; Maier et al., 2022; Pierce et al., 2021; Robertson et al., 2016; Wood-Downie et al., 2021).

Studies using MEGA-PRESS had significant negative group effect size (Hedges'  $g = -0.52$ , 95% CI:  $-0.92 - -0.123$ ,  $I^2 = 83$ ,  $n = 17$ ,  $p < 0.05$ ; across all metabolites and brain regions; Figure

14). No significant effect size was observed for HERMES, STEAM, MEGA-SPECIAL, semi-LAZER, or JPRESS; Figure 14. Note 100% of studies using MEGA-PRESS were observing GABA.

When data were grouped by metabolite, for GABA (supplementary figure 4), studies that used MEGA-PRESS had a significant negative effect size (Hedges'  $g = -0.52$ , 95% CI:  $-0.92 - -0.12$ ,  $I^2 = 83$ ,  $n = 17$ ,  $p < 0.05$ ), which makes sense given the specificity of the MEGA-PRESS sequence. For GABA studies, as  $n = 1$  for STEAM, MEGA-SPECIAL and HERMES acquisitions, results are excluded from interpretation.

For studies that used PRESS to measure NAA, a significant negative group effect size was observed (Hedges'  $g = -0.24$ , 95% CI:  $-0.45 - -0.024$ ,  $I^2 = 74$ ,  $n = 25$ ,  $p < 0.05$ ; supplementary figure 5). A non-significant group effect size was observed for studies measuring NAA using STEAM (Hedges'  $g = -0.66$ , 95% CI:  $-1.46 - 0.13$ ,  $I^2 = 86$ ,  $n = 9$ ,  $p = 0.091$ ). One study used MEGA-SPECIAL and one PROBE-P for NAA measurement (supplementary figure 5).

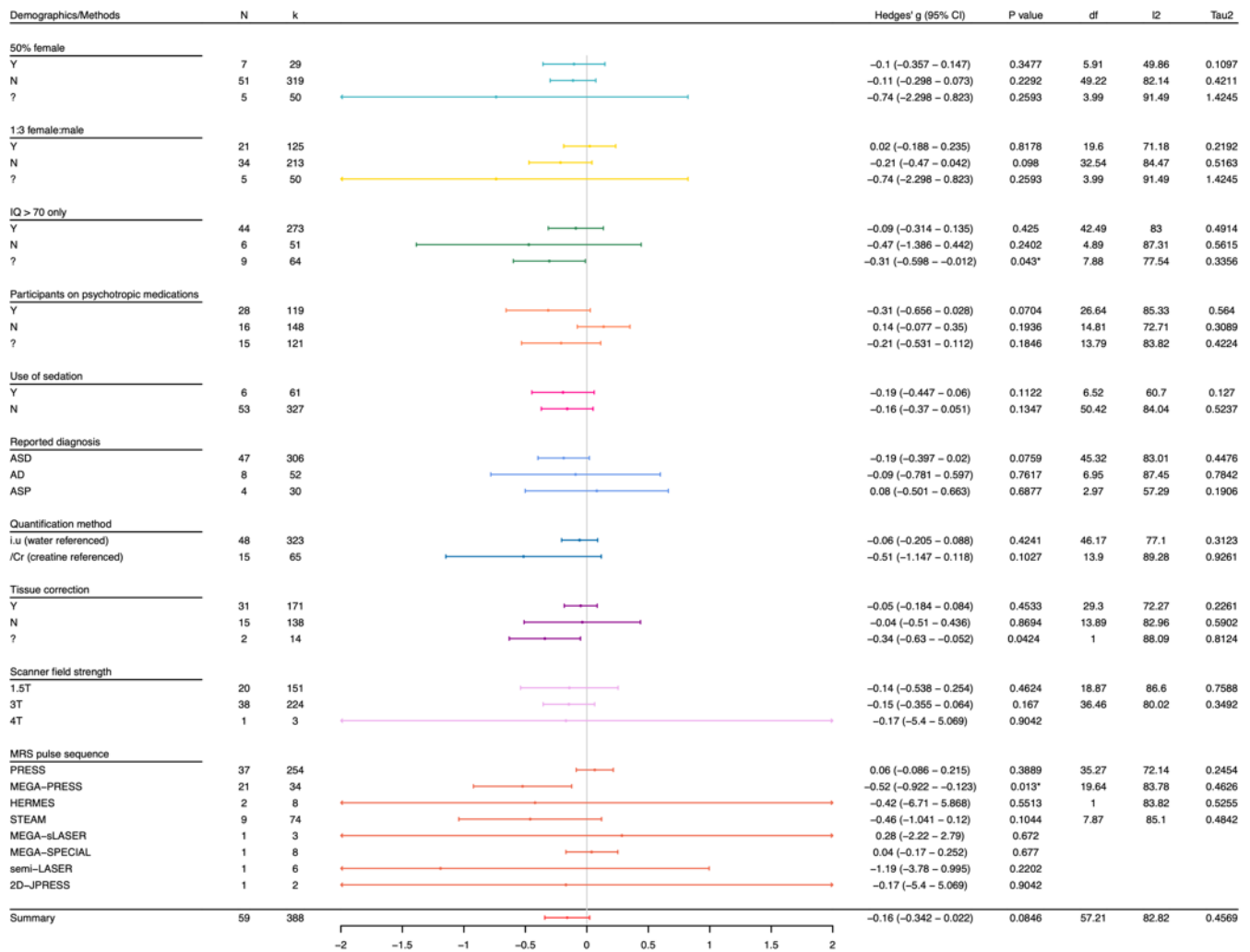
### **Voxel size and transients**

Voxel size and the number of transients showed no relationship with study effect size across all metabolites (Spearman's rho correlation =  $-0.019$ ,  $p = 0.72$  and Spearman's rho correlation =  $0.1$ ,  $p = 0.0902$  respectively; Figure 15).

Data were split by metabolite and the correlation analysis was repeated. The number of transients had a significant positive correlation with effect size for NAA (Spearman's rho correlation =  $0.43$ ,  $p = 0.000$ ) and creatine (Spearman's rho correlation =  $0.47$ ,  $p = 0.000$ ), and ml (Spearman's rho correlation =  $0.42$ ,  $p = 0.044$ ). No significant correlation was

observed between the number of transients and effect size for GABA (Spearman's rho correlation = 0.22,  $p = 0.120$ ), Glx (Spearman's rho correlation = 0.14,  $p = 0.377$ ), Glu (Spearman's rho correlation = -0.43,  $p = 0.0534$ ), GSH (Spearman's rho correlation = -0.7,  $p = 0.122$ ) and choline (Spearman's rho correlation = -0.7,  $p = 0.122$ ).

There was no significant correlation between voxel size and effect size for any metabolite; GABA (Spearman's rho correlation = -0.24,  $p = 0.115$ ), Glu (Spearman's rho correlation = -0.25,  $p = 0.214$ ), GSH (Spearman's rho correlation = -0.49,  $p = 0.325$ ); Glx (Spearman's rho correlation = 0.039,  $p = 0.771$ ), NAA (Spearman's rho correlation = 0.08,  $p = 0.476$ ), choline (Spearman's rho correlation = -0.018,  $p = 0.891$ ), creatine (Spearman's rho correlation = 0.081,  $p = 0.546$ ) and ml (Spearman's rho correlation = 0.14,  $p = 0.388$ ).



**Figure 14.** Forest plot summary of data grouped by demographic and MRS acquisition

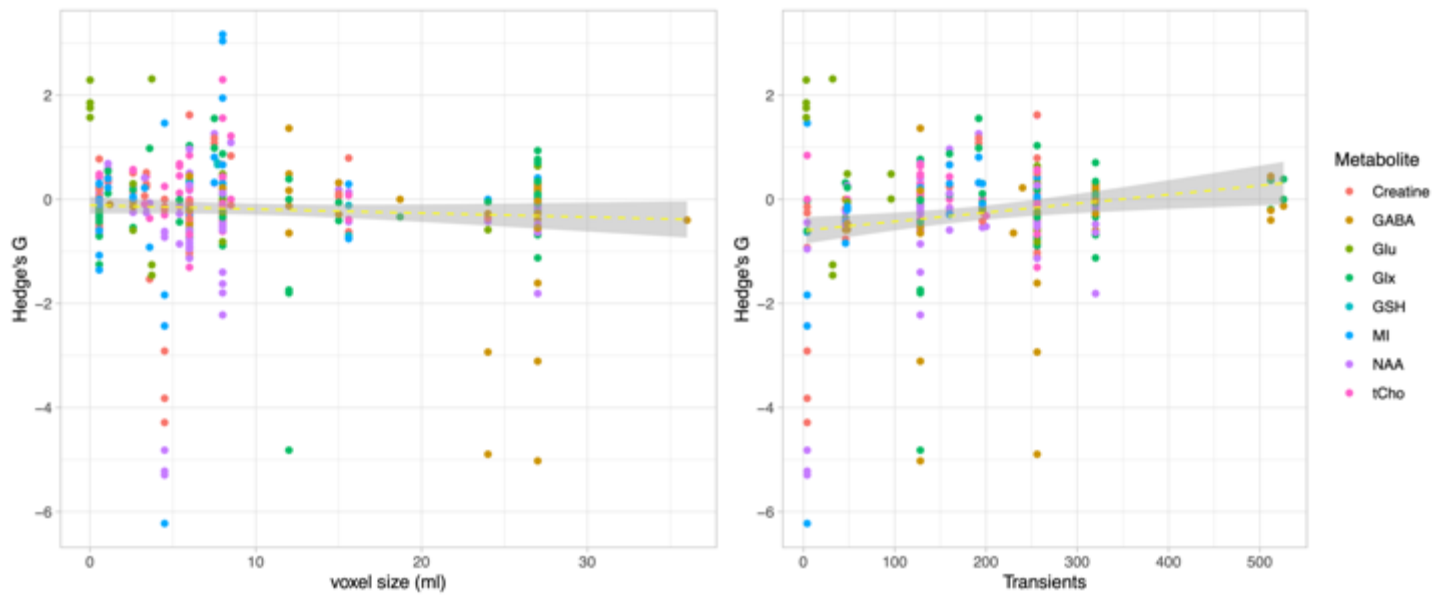
parameters. N: number of studies, k: number of observations, % children: percentage of

studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance

in true effect sizes (another measure of between study heterogeneity). Hedge's G is

reported in respect to autism. \*Statistically significant at p < 0.05, and at p < 0.01 when the

degrees of freedom < 4 for RVE t-tests.



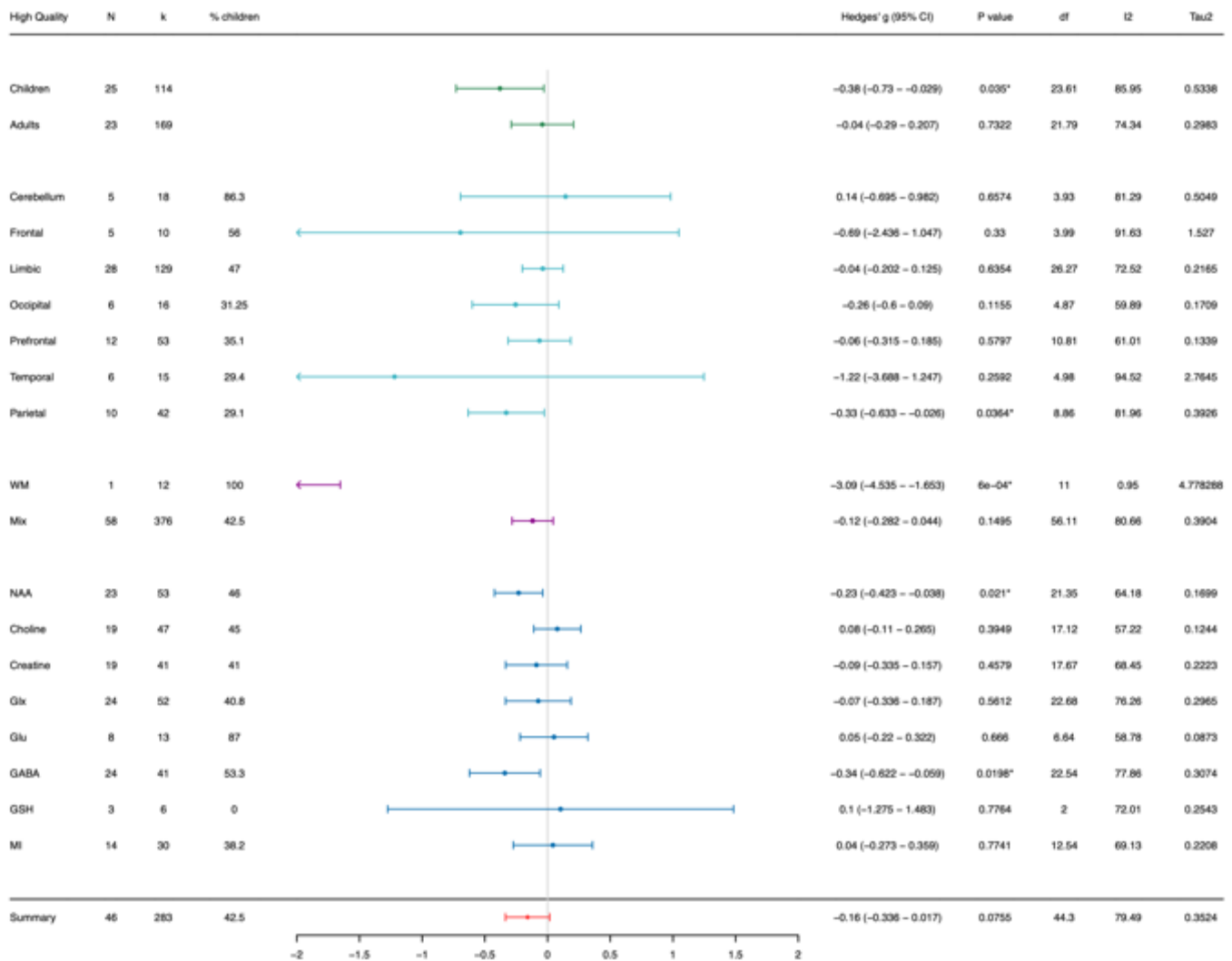
**Figure 15. Relationship between study effect size and voxel size/number of transients.**

Voxel size (ml) showed no relationship with study effect size (Spearman's rho correlation = -0.019,  $p = 0.72$ ), while the number of transients showed a positive correlation with study effect size (Spearman's rho correlation = 0.1,  $p = 0.0902$ ).

### 3.9 Effect of MRS data quality

Analysis was repeated using only high-quality data as assessed by MRS -Q (Figure 16).

General trends identified in the main meta-analysis were conserved, with small changes to effect size values but with no change in significance level. Across all brain regions and ages, significant negative group effect sizes were observed for NAA and GABA (Hedges'  $g_{\text{NAA}} = -0.23$ , 95% CI: -0.42 - -0.038,  $I^2 = 64$ ,  $n = 23$ ,  $p < 0.05$  and Hedges'  $g_{\text{GABA}} = -0.34$ , 95% CI: -0.62 - -0.059,  $I^2 = 77$ ,  $n = 25$ ,  $p < 0.05$ ).



**Figure 16.** Data from only 'high quality' studies as assessed by MRS-Q, grouped by age (across all metabolites, brain regions and ages), brain regions (across all metabolites and ages), voxel tissue type (across all metabolites and ages), and metabolite of interest (across all brain regions, ages, and tissue types). N: number of studies, k: number of observations, % children: percentage of studies observing children only, I<sup>2</sup>: measure of between study heterogeneity, Tau<sup>2</sup>: Variance in true effect sizes (another measure of between study

heterogeneity). \*Statistically significant at  $p < 0.05$ , and at  $p < 0.01$  when the degrees of freedom  $< 4$ .

## 4. Discussion

### Summary of findings across all regions and ages

Here we provide evidence of significant alterations in brain metabolite concentrations in autism. Firstly, across all metabolites, we observe significantly lower metabolite concentrations in autistic children compared to non-autistic controls, whilst we observe no significant neurochemical differences between groups in adults. These findings support other MRI work suggesting that autistic brain differences are more pronounced in childhood compared to adulthood, potentially due to compensatory mechanisms (Courchesne, 2004; Schumann et al., 2004). However, we do find that these results are driven by changes in specific metabolites/brain regions which we explore at a mechanistic level below.

### Autism-associated alterations in NAA.

We show that limbic and prefrontal NAA concentrations are significantly lower in autistic children compared to non-autistic controls. Of the limbic regions, this trend was potentially driven by the **anterior cingulate cortex (ACC)**, however as this region was considerably more statistically powered than other limbic areas, this should be interpreted with caution.



Autism-associated differences in ACC and prefrontal NAA concentrations may indicate differences in neuronal integrity within these brain regions, NAA being highly concentrated within the neuronal cell body (Rae, 2014). Consistent with this, the immature autistic brain displays global and regional structural differences when compared to non-autistic controls (Courchesne, 2004; Courchesne et al., 2001; Hazlett et al., 2005), including reductions in ACC and prefrontal GM and WM volumes as documented by various studies (Chien et al., 2021; Haznedar et al., 1997; Knaus et al., 2012; Yang et al., 2016), as well as reductions in ACC neuronal size and dendritic spine complexity (Kemper & Bauman, 1993; Simms et al., 2009). Lower NAA concentrations in childhood may also indicate untypicalities in developmental myelination, with decreased NAA concentrations in oligodendrocytes previously associated with myelin sheath lipid content abnormalities (Singhal et al., 2017). This is consistent with autism-associated reductions in the fractional anisotropy of ACC and prefrontal white matter tracts (indicative of demyelination and/or reduced nerve fibre coherence; [Barnea-Goraly et al., 2004](#); [Basser, 1995](#); [Beaulieu, 2002](#); [Hau et al., 2019](#); [Keller et al., 2007](#)) and post-mortem work finding autism associated reductions in prefrontal axon myelin thickness (Zikopoulos & Barbas, 2010). Finally, as NAA is synthesised in mitochondria (Siddiqui et al., 2016), differences in mitochondrial function in these regions may also contribute, with mitochondrial dysfunction recently linked to many autism aetiologies (see review; Siddiqui et al., 2016). These are not discrete factors; atypical mitochondrial function alters neuronal differentiation and connectivity (various models: Fernandez et al., 2019; Khacho et al., 2017; Klein Gunnewiek et al., 2020) and as such may contribute to the differences in neuronal integrity during brain development (Fernandez et al., 2019; Siddiqui et al., 2016).

These structural and cytoarchitectural differences may underlie observed differences in the functional activity and connectivity of both the prefrontal cortex (Cole et al., 2019; Gilbert et al., 2008; N. Schmitz et al., 2006; Woodward et al., 2017) and limbic brain regions (Etherton et al., 2011; Guo et al., 2019; Kana et al., 2007; Turner et al., 2006; Woodward et al., 2017; Y. Zhou et al., 2016) in autism. In turn, functional differences in these regions likely underpin behavioural and cognitive differences in autism, as for example the prefrontal cortex is associated with social cognition, executive function and language (Franklin et al., 2017; Friedman & Robbins, 2022; Mohapatra & Wagner, 2023; Sumiya et al., 2020), while limbic brain regions are involved in several brain processes relevant to autism phenotypes. The ACC, for example, has roles in attention and social cognition (Apps et al., 2016; Caruana et al., 2017; Chien et al., 2021; Dichter et al., 2009). The amygdala has been associated with facial recognition, facial expression processing, and emotional learning (Adolphs et al., 1994; Adolphs, 2008; Rutishauser et al., 2015; Wang et al., 2017). The thalamic nuclei relay primary sensory and motor information from the periphery to the appropriate cortical regions (Lopez-Bendito, 2018), while the hippocampus has been implicated in episodic memory (Scoville & Milner, 1957; Eldridge et al., 2005) and more recently social interaction (Tavares et al., 2015; Banker et al., 2021). As such, researchers find a negative correlation between limbic (ACC) and prefrontal NAA concentrations and autism clinical scores (incl. communication impairments, IQ, autistic mannerisms & sensory processing abnormalities; Gabis et al., 2008; Hegarty et al., 2017; Carvalho Pereira et al., 2017; Brix et al., 2015; Edmondson et al., 2020). NAA differences are therefore mechanistically relevant as, in the context of the wider structural and functional implications, they contribute to (and may help to predict) autism phenotypes and their severity (Denisova & Lin, 2023).

By adulthood, there is a convergence of limbic and prefrontal NAA concentrations between autism and non-autistic controls, with significant differences between the two groups in prefrontal and limbic (ACC) NAA concentrations no longer seen, this finding unbiased by statistical power (same number of studies per age group). As discussed, this is not the first time a ‘normalisation’ of autism brain differences has been identified in the MRI imaging field; significant atypicality’s in brain structure observed during early childhood have been found to be absent in adult cohorts, potentially explaining the loss of NAA differences (Courchesne, 2004; Courchesne et al., 2001; Hazlett et al., 2005). However, this is largely based on cross sectional data (as is our meta-analysis) and is in conflict with other MRI and human post mortem findings that show cytoarchitectural differences at the neuronal and regional level persist in adulthood (Kemper & Bauman, 1993; Simms et al., 2009; Yang et al., 2016). As such, in the absence of longitudinal evidence, a ‘normalisation’ hypothesis should be considered with caution (Yankowitz et al., 2020).

#### Autism-associated differences in GABA and Glutamate (Glu and Glx).

We also find autism-associated differences in GABA concentrations, with **limbic** GABA concentrations significantly lower in autistic compared to non-autistic children. This trend is notably driven by the **ACC** within the limbic regions studied. For autistic adults, we find no significant alterations in GABA concentrations. However, due to the limited number of studies focusing on GABA concentrations in adults, particularly in the ACC/limbic regions, an observation bias may underpin this apparent age effect.

The MRS GABA+ signal is a proxy measure for the degree a circuit can employ inhibition (Rae et al., 2014; Puts et al., 2015). In this context, the observed alterations in GABA concentrations

(given that we observe no significant alterations in Glx/Glu concentrations) are consistent with the E/I imbalance theory of autism (Rubenstein & Merzenich, 2003), specifically showing an uncompensated shift in E/I balance towards excitation. While the relatively poor spatial resolution of the MRS technique means we cannot specify the cellular mechanisms that underpin these observations, previous work observes autism-associated reductions in the expression of the enzymes responsible for GABA synthesis (glutamate decarboxylase 65 and 67; Fatemi et al., 2002), as well as autism-associated decreases in parvalbumin (PV+) interneuron density, the primary cortical inhibitory (and so GABA expressing) interneuron (Hashemi et al., 2017). Both could sensibly contribute to the observed differences in limbic (ACC) GABA concentrations.

At the mechanistic level, local GABA concentrations have previously been associated with functional magnetic resonance imaging (fMRI) task response amplitudes (Muthukumaraswamy et al., 2009; Sampaio-Baptista et al., 2015; T. W. Schmitz et al., 2017) and due to GABAergic regulation of neuronal synchrony within local and regional brain circuits (Baumgarten et al., 2016; Muthukumaraswamy et al., 2009), the functional connectivity strength of various networks (Muthukumaraswamy et al., 2009; Stagg et al., 2014; Assaf et al., 2010; Sampaio-Baptista et al., 2015; Y. Zhou et al., 2016). Both ACC functional connectivity and task-induced activation have been shown to be atypical in autism (Agam et al., 2010; Assaf et al., 2010; Cole et al., 2019; Kana et al., 2007; Y. Zhou et al., 2016) and have been associated with repetitive behaviours and social differences (Chien et al., 2021; Thakkar et al., 2008; Yang et al., 2016). This provides a mechanistic link between differences in limbic (ACC) GABAergic activity and autism phenotypes, with previous work finding negative correlations between ACC GABA concentrations and autism clinical scores (Carvalho Pereira et al., 2017;

Edmondson et al., 2020; Brix et al., 2015). Our results may thus help to explain why pharmacological potentiators of the GABAergic system improve behavioural and cognitive phenotypes in mouse models of autism (Baclofen; Silverman et al., 2015, Gaboxadol; Cogram et al., 2019) and in autistic children (Bumetanide; Lemonnier et al., 2012, 2017; Mahdaviniasab et al., 2019; Zhang et al., 2020). However, according to our results, shifts in E/I balance may not be global or equal across the brain, whilst GABAergic drugs act without spatial specificity and so may have (counter-productive) off target effects. This potentially explains the null findings during phase 2 autism clinical trials for Arbaclofen, a GABA-b receptor agonist (and so potentiator; Veenstra-VanderWeele et al., 2017), as well as the sometimes moderate side effects reported with other agents (Bumetanide; Zhang et al., 2020).

There are several limitations specific to the measurement of GABA and Glu/Glx that require some discussion. First, most of studies included in this meta-analysis measured GABA concentrations only, yet decreases in GABA may lead to compensatory decreases in Glu (Kubas et al., 2012). While we observe no significant differences in Glu/Glx concentrations across all brain regions and ages (Glu/Glx observations were of a similar number to GABA), to resolve this definitively, multiple measurements should be taken within the same cohort or even at the individual level. Furthermore, care must be taken to equate MRS-measured GABA and Glu to neurotransmission E/I balance given their role in energy metabolism (Rae 2014). These metabolites are in constant flux as GABA (and GSH) are synthesised downstream of Glutamate/Glutamine. As such, MRS GABA/Glu balance may only be an indicator of broader tonic E/I in the brain.

Secondly, roughly half of studies measuring GABA reported concentrations as GABA plus macromolecule (MM) contamination of signal (GABA+), while the remaining studies used MM modelling to obtain pure GABA estimates. While we did not report on the effect of this on study outcome (as modelling approaches varied between studies and thus were likely not comparable) and the role of MM in autism is likely to be relatively small, it does represent another source of potential bias. Although MM-suppressed editing sequences are available, they are particularly sensitive to motion due to subtraction of non-interlaced transients (Edden et al., 2016) and as such may not be optimal in populations who tend to move more in the scanner environment such as those with cognitive difficulties or autism (Yerys et al., 2009, Puts et al., 2017).

Lastly, we recorded Glu and Glx separately and findings were consistent. However, it remains unclear whether Glu can be measured precisely at 1.5T and 3T and can be measured separately from Glx (glutamate + glutamine). In addition, quantification of Glx will differ between studies, some having quantified this from unedited MRS, others from the Glx peak in the GABA-edited difference spectrum, and others from the SUM (edit-ON + edit-OFF) or edit-OFF spectrum. It is well-known that measures derived using these different approaches do not always equate to one another (Bell et al., 2021). While Glu appears to contribute in majority to MRS Glx (as opposed to Gln; F. Du et al., 2009; Tapper et al., 2020), further caution should be taken interpreting measurements in the context of E/I balance.

### Choline, Creatine and ml.

The number of studies (and percentage of child cohorts) observing creatine, choline and ml was not dramatically smaller than those observing GABA and NAA. As such, we do not think

that our lack of significant findings for these metabolites is due to a bias in the total number of observations. There is however some evidence of autism-associated differences in these metabolites. We find significantly lower parietal choline (and NAA) concentrations and cerebellar creatine concentrations in autism. These changes were not significant in either the children only or adults only subgroups (only when combined across all ages), likely due to a lack of power from the smaller study numbers.

MRS total choline is composed of the choline containing compounds phosphocholine, phosphatidylcholine and glycerophosphocholine (Miller et al., 1996; Rae, 2014), which together are essential for neuronal membrane integrity (Rae, 2014). Reduced parietal total choline (and NAA) may therefore be indicative of previously characterised structural differences in this region, including reductions in GM and parietal cortical thickness (Ford & Crewther, 2016b; McAlonan et al., 2005; Wallace et al., 2010).

MRS creatine signal is composed of creatine, which when reported as total creatine includes its phosphorylated form phosphocreatine, which acts as a readily available and fast energy store within neurons (Rae, 2014; Wallimann et al., 1992). The cerebellum is widely understudied in autism, however differences in cerebellar creatine may reflect abnormalities in neuronal activity within this region, in keeping with previously described alterations in cerebral-cerebellar functional connectivity and activity in autistic individuals (Ford & Crewther, 2016a; Mapelli et al., 2022; Mostofsky et al., 2009).

### Redox hypothesis (GSH)

Recent advances in MRS acquisition sequences have facilitated the quantification of GSH *in vivo* (Terpstra et al., 2003) through spectral editing, and so may provide evidence to support the redox hypothesis. Of the six relevant studies, no significant differences in GSH concentrations between autistic adults and non-autistic controls were identified (Limbic regions, frontal regions and occipital regions). If GSH conforms with the general trend identified in this meta-analysis, future work may find more significant GSH differences in childhood.

### Effect of MRS study design

Demographic and methodological factors likely explain why between study heterogeneity ( $I^2$ ) was large, even when data were simultaneously grouped by metabolite, age and brain region. As such, we explored how these factors associated with effect size.

#### *Sex*

Males are diagnosed with autism more often than females, with a reported 1:3 ratio of females to males in autistic populations (Loomes et al., 2017). Recent work suggests that the presentation of autism in females may simply not be captured with current gold standard tools used by researchers to confirm autism diagnosis, leading to underestimates of female autism prevalence and the exclusion of females from research studies at significantly greater rates than males (de Giambattista et al., 2021; D’Mello et al., 2022). In this meta-analysis, only 16% of studies included cohorts matching population prevalence (1:3 female:male), only 11% included sex-matched cohorts (1:1 female:male), while 22% of included studies excluded females altogether. Failure to include sex matched samples



diminishes statistical power for sex comparison, which is vital as the underlying biochemistry of autism may differ between sexes (Fung et al., 2021). As an example, we observed lower concentrations of NAA and GABA in autism *only* when female:male ratios of cohorts were less than 1:3/1:1. Thus we cannot be certain that the main outcomes of this meta-analysis are not specifically biased for male populations, with potentially subtler metabolite differences in autistic females. Future MRS study design should thus enable reliable statistical evaluation of sex differences to address this.

### *IQ*

75% of studies included in this meta-analysis excluded individuals with an IQ < 70, although it is estimated that around 30% of autistic children have an intellectual disability (Christensen et al., 2016; Zablotsky, 2017). As such, our results are more applicable to individuals without intellectual disability (i.e., having low support needs). Across all metabolites, our findings show tendencies that neurometabolic alterations in individuals with IQ > 70 are subtler than those with more severe intellectual impairments. This agrees with studies showing that altered neurometabolite levels (specifically NAA and GABA) linearly correlate with IQ (Carvalho Pereira et al., 2018; Gabis et al., 2008; Hegarty et al., 2018). While IQ should thus be considered as a potentially confounding factor in MRS studies of autism, discounting those with intellectual disabilities reduces the generalizability of study results, an important consideration as those most likely to be eligible/desire therapeutic intervention are those with higher support needs. Thus, excluding those with IQ < 70 with the sole aim of controlling for IQ is not necessarily an advantageous or suitable approach if justifying the same study with therapeutic goals.

### *Sedation and psychotropics drugs*

According to a study in 2017, 50% of autistic children take psychotropic medication (Madden et al., 2017). Here, 47% of included studies contained participants using psychotropic medications (n = 28), with SSRI's being the most common. Psychotropic medication use by autistic individuals may indicate more severe symptom profiles, as well as the treatment of co-occurring psychiatric conditions such as anxiety disorder, ADHD and depression, which are very common in autism (72% of autistic individuals; Strang et al., 2012; Hours et al., 2022)) and also associated with alterations in the GABAergic system (Edden et al., 2012; Hasler et al., 2007; Pollack et al., 2008). Thus, the direct impact of psychotropic medications on brain metabolites (Sanacora et al., 2002) in addition to the presence of co-occurring disorders and/or an increased phenotype severity may contribute to our findings that GABA alterations are more pronounced in studies including participants on psychotropic medication compared to studies excluding such participants. As only 24% of the studies included in this meta-analysis controlled (statistically) for psychotropic medication usage in their analysis and only half reported co-occurring conditions. In future, study designs should enable statistical comparison of medicated and non-medicated groups, and transparent reporting of co-occurring disorders is recommended.

10% of the included studies used sedation during MRI procedures, most commonly using triclofos sodium. Across studies, there was considerable variability of both the dose and proportion of sedated participants (from less than half – all). Based on these studies alone, we observed no clear effect of sedation on study outcome. While the impact of triclofos sodium on brain metabolites has not been systematically assessed, these findings are consistent with a number of studies that find no impact of sedation or anaesthesia on brain

metabolite NAA, creatine and Glu concentrations (F. Du et al., 2009; pentobarbital; Jacob et al., 2012; Sevoflurane and Propofol) . Note however that the effect of anaesthesia on GABA levels in particular is undefined.

### *Diagnosis*

In line with current DSM-5 guidelines, studies observing AD, AS, and ASD were included in this meta-analysis. Nevertheless, across all metabolites, we observed trends for greater metabolite differences in ASD cohorts compared to those classified to either an AS and/or AD. More specifically, significantly lower GABA concentrations were observed in ASD cohorts only, whilst significantly lower NAA concentrations were observed in AD and ASD cohorts only. As such, the main outcomes of this meta-analysis are at this stage applicable to ASD cohorts (reported by 72% of included studies), partially applicable to AD cohorts (reported by 13% of studies included) and not applicable to AS cohorts (reported by only 6% of studies included). As AS diagnosis historically included mainly high-functioning individuals with ‘milder’ phenotypes compared to AD (Colleen & Stone, 2014), it is perhaps not surprising that no significant metabolite differences are observed for this group, particularly given that metabolite alterations correlate with phenotype severity (see above). In future, exploring the neurochemistry of non-subjective clinical behaviours maybe a preferable approach to mitigate the confounding nature of evolving diagnostic criteria between studies.

### *Quantification method*

We observed trends for stronger autism-associated neurometabolite differences in creatine referenced data compared to water referenced. This is striking as differences in GABA concentrations were only significant when creatine referenced, yet 60% of studies reported

water-referenced GABA concentrations. There are multiple factors that may contribute to these results. First, creatine referencing typically assumes that creatine does not differ in the autistic brain which may not be the case; as evidenced by the observed significant differences in cerebellum creatine concentrations between autism and non-autistic groups. On the other hand, water referencing requires corrections for tissue composition and metabolite relaxation to obtain estimated metabolite concentrations (as recommended by MRS-Q). Given the breadth of studies on anatomical differences in autism at both the micro- and meso-level (see review; (Donovan & Basson, 2017), this complicates partial volume correction and introduces various unknowns. Both methods thus have their limitations and so it would be beneficial to study the structural and neurochemical associations in autism in more detail. However, at present, it is suggested that researchers report both ratio and estimated concentrations for transparency purposes.

### *Scanner field strength*

Although we observed no effect of scanner field strength on study outcome overall, for NAA we observed significantly lower concentrations in autism at 1.5T and not 3T (although the effect size at 3T remained negative). Since 3T provides a higher SNR compared to 1.5T, this may be cause for concern (Di Costanzo et al., 2007; Wilson et al., 2019, 2019). However, NAA is one of the most highly concentrated metabolites on the MRS chemical shift spectra/axis, making it one of the most accurate to quantify at low field strengths (1.5T) (Peek et al., 2020). As such, we can be confident that the significance of findings at 1.5T is not just a result of extra noise (the same could not be said for low concentration metabolites for example Glx/GABA). Furthermore, the significance of NAA findings was

conserved when low quality studies were excluded (discussed below), providing further reassurance that the observed data trends are not skewed/inflated by low quality data.

### *Pulse sequence*

The impact of MRS pulse sequence on study effect size is more challenging to assess since study number per sequence was often too small to comment ( $n = 1$  for MEGA-sLASER,, MEGA-SPECIAL, semi-LASER, 2D-JPRESS). Across all metabolites, we observed significant alterations for studies using MEGA-PRESS editing sequence only. As 100% of studies using MEGA-PRESS observed GABA concentrations, this was likely driven by the observed reductions in GABA concentrations in autism as opposed to a bias introduced by the sequence itself. Lower NAA concentrations (in autism) were observed by studies using PRESS and STEAM non-editing sequences, while this was only significant for PRESS. Due to differences in sequence design, STEAM achieves more accurate volume selection (less chemical shift displacement) compared to PRESS, however at the expense of SNR, achieving around half the SNR of PRESS (van der Graaf, 2010) . As such, perhaps with greater SNR (using PRESS) we can better detect NAA differences in autism. The discussion below however challenges this notion for NAA.

### *Transients and voxel size*

Increasing the number of transients and voxel size increases the accuracy of MRS metabolite quantification (due to increasing signal to noise ratio (SNR); Mikkelsen et al., 2018).

Optimally, increasing accuracy should increase the ability to detect differences in metabolite concentrations between groups if present (as suggested above). As such it was surprising that overall, we observed no significant interaction between voxel size/transients and study

effect size. In fact, for NAA, we found a significant positive correlation between the number of transients and study effect size. This suggests that with increasing accuracy studies can better detect increases in NAA concentrations in autism (hedges  $g$  is positive), and as such queries the validity of our main results. However, multiple MRS parameters (alongside transients) contribute to MRS data quality, as was comprehensively assessed by MRS-Q. MRS-Q is thus perhaps a better assessment of whether the observed data trends are reliable than looking at the number of transients alone.

### *Publication bias*

For studies observing GABA and NAA we observed a significant publication bias, with study effect sizes positively skewed (absence of effect sizes in the positive direction). Note that all of the GABA studies and the majority of the NAA that studies were identified as driving this asymmetry (large, significant negative effect sizes) reported creatine scaled metabolite concentrations (Gabis et al., 2008; Gaetz et al., 2014; Kubas et al., 2012; Port et al., 2019), reiterating the potential bias quantification method introduces to study outcome.

Importantly however, between study heterogeneity was large and so included studies likely do not truly share one true effect size (due to real differences in demographics/methodology), making such bias assessments potentially inappropriate (Peters et al. 2007; Terrin et al. 2003; Simonsohn et al., 2014). As such the identified studies remained in our meta-analysis.

Biases out of the scope of a meta-analysis do exist, as identified in our AXIS assessment; the use of hospital controls instead of the general population (often migraine patients), failure to characterises individuals who failed to complete the MRS scan, failing to control for co-

occurring conditions such as ADHD and anxiety in particular, and variation in individual researcher MRS processing pipelines (Bhogal et al., 2017) all likely affect these data. Elimination of all confounding factors in autism case-control studies, this being a highly heterogenous and complex group, is perhaps unrealistic. Better and transparent reporting practises would make statistical control for this bias easier and interpretation of results more reliable.

### *Effect of MRS quality on study outcome*

The development of consensus guidelines (MRSinMRS and MRS-Q; A. Lin et al., 2021; Peek et al., 2020) is essential for improvement of adequate reporting and acquisition of MRS data. Without this, it remains a considerable challenge to isolate non-confounded data trends as proven by this meta-analysis. This consensus was used to assess the quality of MRS study design and data. Low quality studies (10%) generally did not report sufficient scanner field strength, voxel size and transients (consensus edited; 3T and 27 ml voxel for edited MRS for GABA, 240 transients. Unedited; 3.4 ml voxel, 128 transients at 3T). As discussed, these parameters are associated with SNR and directly influence the quality of the acquired MRS spectra (Mikkelsen et al., 2018). To prevent the skewing of data trends due to poor quality data, we excluded these studies from analysis. All significant trends from the main meta-analysis were conserved; significantly lower NAA and GABA concentrations in autism compared to TD. As such, we present high-quality evidence of autism-associated alterations for NAA and GABA concentrations in limbic (and parietal) regions. While MRS-Q addresses MRS methodological factors, these results are still subject to the demographic confounds discussed above.

## Limitations of the current meta-analysis

One limitation to this meta-analysis is that we initially included all studies regardless of quality as assessed by MRS-Q. However, we observed no changes to the data trends when low quality data were excluded, and we used subgrouping to thoroughly explore potential biases in study outcomes. We have also been very transparent in reporting of these (MRS study design and demographics). For some brain regions we lacked statistical power due to the small number of studies (particularly the temporal and occipital regions), and as such the lack of significant differences in neurometabolite concentrations between groups should be interpreted with caution as it may simply reflect that we are unable to detect differences at this time.

A key strength of the current analysis is that we employed a conservative RVE method for Hedges'  $g$  calculation (Pustejovsky & Tipton, 2022), this prevented the skewing of outcomes by dependent data, with most studies taking multiple metabolite measurements (and voxels) from one individual (Borenstein et al., 2021). This was not considered by previous systematic reviews (Ajram et al., 2019; Y. Du et al., 2023; Ford & Crewther, 2016a). We have also comprehensively assessed the impact of study design on outcome to expose confounding factors in our results.

## Conclusion

Overall, in this meta-analysis, we find evidence of strong autism-associated alterations in neuro-metabolite concentrations. More specifically, we find that autistic *children* have significantly lower concentrations of NAA in the limbic and prefrontal regions, perhaps



indicative of structural differences and/or mitochondrial dysfunction in this region. Also reported was evidence of uncompensated shifts in E/I imbalance; autistic *children* have significantly lower GABA concentrations in limbic regions (more specifically the ACC), while we observe no significant alteration in Glx/glutamate concentrations. The severity of alterations in both NAA and GABA has been shown to correlate with the severity of autism clinical scores, supporting the hypothesis that such atypicality's contribute mechanistically to the autism phenotype, with GABA specifically representing a valid therapeutic target. Finally, we observed that sex, co-occurring disorders, psychotropic medication use and MRS metabolite quantification approach influence the degree of GABA/NAA alteration observed and so likely contribute to the heterogeneity of previous findings. Compliance with consensus MRS parameters (e.g MRS-Q), paired with appropriate and transparent reporting is essential to address this and improve the ease at which reliable trends can be identified and accepted in future.

## 5. Acknowledgements

TA was supported by an MRC Clinician Scientist Fellowship [MR/P008712/1] and Transition Support Award [MR/V036874/1]. AT, NP and TA received support from the Medical Research Council Centre for Neurodevelopmental Disorders, King's College London [MR/N026063/1]. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

## 6. Declarations of interest

Authors have no declarations of interest to disclose.

## 7. Data availability

Data are available through: [link]

## 8. Bibliography

- Adams, J. B., Baral, M., Geis, E., Mitchell, J., Ingram, J., Hensley, A., Zappia, I., Newmark, S., Gehn, E., Rubin, R. A., Mitchell, K., Bradstreet, J., & El-Dahr, J. M. (2009). The Severity of Autism Is Associated with Toxic Metal Body Burden and Red Blood Cell Glutathione Levels. *Journal of Toxicology*, 2009, e532640. <https://doi.org/10.1155/2009/532640>
- Agam, Y., Joseph, R. M., Barton, J. J. S., & Manoach, D. S. (2010). Reduced cognitive control of response inhibition by the anterior cingulate cortex in autism spectrum disorders. *NeuroImage*, 52(1), 336–347. <https://doi.org/10.1016/j.neuroimage.2010.04.010>
- Ajram, L. A., Pereira, A. C., Durieux, A. M. S., Velthuis, H. E., Petrinovic, M. M., & McAlonan, G. M. (2019). The contribution of [1H] magnetic resonance spectroscopy to the study of excitation-inhibition in autism. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 89, 236–244. <https://doi.org/10.1016/j.pnpbp.2018.09.010>
- American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR)* (4th ed., Vol. 1). American Psychiatric Association. <https://doi.org/10.1176/appi.books.9780890423349>
- Aoki, Y., Abe, O., Yahata, N., Kuwabara, H., Natsubori, T., Iwashiro, N., Takano, Y., Inoue, H., Kawakubo, Y., Gono, W., Sasaki, H., Murakami, M., Katsura, M., Nippashi, Y., Takao, H., Kunimatsu, A., Matsuzaki, H., Tsuchiya, K. J., Kato, N., ... Yamasue, H. (2012).

- Absence of age-related prefrontal NAA change in adults with autism spectrum disorders. *Translational Psychiatry*, 2(10), e178. <https://doi.org/10.1038/tp.2012.108>
- Apps, M. A. J., Rushworth, M. F. S., & Chang, S. W. C. (2016). The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. *Neuron*, 90(4), 692–707. <https://doi.org/10.1016/j.neuron.2016.04.018>
- Assaf, M., Jagannathan, K., Calhoun, V. D., Miller, L., Stevens, M. C., Sahl, R., O’Boyle, J. G., Schultz, R. T., & Pearlson, G. D. (2010). Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients. *NeuroImage*, 53(1), 247–256. <https://doi.org/10.1016/j.neuroimage.2010.05.067>
- Balduzzi, S., Rücker, G., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence Based Mental Health*, 22(4), 153–160. <https://doi.org/10.1136/ebmental-2019-300117>
- Barnea-Goraly, N., Kwon, H., Menon, V., Eliez, S., Lotspeich, L., & Reiss, A. L. (2004). White matter structure in autism: Preliminary evidence from diffusion tensor imaging. *Biological Psychiatry*, 55(3), 323–326. <https://doi.org/10.1016/j.biopsych.2003.10.022>
- Basser, P. J. (1995). Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine*, 8(7–8), 333–344. <https://doi.org/10.1002/nbm.1940080707>
- Baumgarten, T. J., Oeltzschner, G., Hoogenboom, N., Wittsack, H.-J., Schnitzler, A., & Lange, J. (2016). Beta Peak Frequencies at Rest Correlate with Endogenous GABA+/Cr Concentrations in Sensorimotor Cortex Areas. *PLOS ONE*, 11(6), e0156829. <https://doi.org/10.1371/journal.pone.0156829>

- Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system—A technical review. *NMR in Biomedicine*, *15*(7–8), 435–455.  
<https://doi.org/10.1002/nbm.782>
- Begg, C. B., & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics*, *50*(4), 1088–1101. <https://doi.org/10.2307/2533446>
- Bejjani, A., O’Neill, J., Kim, J. A., Frew, A. J., Yee, V. W., Ly, R., Kitchen, C., Salamon, N., McCracken, J. T., Toga, A. W., Alger, J. R., & Levitt, J. G. (2012). Elevated Glutamatergic Compounds in Pregenual Anterior Cingulate in Pediatric Autism Spectrum Disorder Demonstrated by 1H MRS and 1H MRSI. *PLOS ONE*, *7*(7), e38786.  
<https://doi.org/10.1371/journal.pone.0038786>
- Bell, T., Boudes, E. S., Loo, R. S., Barker, G. J., Lythgoe, D. J., Edden, R. A., Lebel, R. M., Wilson, M., & Harris, A. D. (2021). IN VIVO GLX AND GLU MEASUREMENTS FROM GABA-EDITED MRS AT 3T. *NMR in Biomedicine*, *34*(5), e4245.  
<https://doi.org/10.1002/nbm.4245>
- Bernardi, S., Anagnostou, E., Shen, J., Kolevzon, A., Buxbaum, J. D., Hollander, E., Hof, P. R., & Fan, J. (2011). In vivo 1H-magnetic resonance spectroscopy study of the attentional networks in autism. *Brain Research*, *1380*, 198–205.  
<https://doi.org/10.1016/j.brainres.2010.12.057>
- Bernardino, I., Dionísio, A., Violante, I. R., Monteiro, R., & Castelo-Branco, M. (2022). Motor Cortex Excitation/Inhibition Imbalance in Young Adults With Autism Spectrum Disorder: A MRS-TMS Approach. *Frontiers in Psychiatry*, *13*, 860448.  
<https://doi.org/10.3389/fpsy.2022.860448>
- Bhogal, A. A., Schür, R. R., Houtepen, L. C., van de Bank, B., Boer, V. O., Marsman, A., Barker, P. B., Scheenen, T. W. J., Wijnen, J. P., Vinkers, C. H., & Klomp, D. W. J. (2017). 1H-MRS

processing parameters affect metabolite quantification: The urgent need for uniform and transparent standardization. *NMR in Biomedicine*, 30(11), e3804.

<https://doi.org/10.1002/nbm.3804>

Bjørklund, G., Tinkov, A. A., Hosnedlová, B., Kizek, R., Ajsuvakova, O. P., Chirumbolo, S., Skalnaya, M. G., Peana, M., Dadar, M., El-Ansary, A., Qasem, H., Adams, J. B., Aaseth, J., & Skalny, A. V. (2020). The role of glutathione redox imbalance in autism spectrum disorder: A review. *Free Radical Biology and Medicine*, 160, 149–162.

<https://doi.org/10.1016/j.freeradbiomed.2020.07.017>

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to Meta-Analysis*. John Wiley & Sons.

Bottomley, P. A. (1987). Spatial Localization in NMR Spectroscopy in Vivo. *Annals of the New York Academy of Sciences*, 508(1), 333–348. <https://doi.org/10.1111/j.1749-6632.1987.tb32915.x>

Bowden, J., Davey Smith, G., & Burgess, S. (2015). Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *International Journal of Epidemiology*, 44(2), 512–525.

<https://doi.org/10.1093/ije/dyv080>

Brian, J. A., Doyle-Thomas, K., Baribeau, D., & Anagnostou, E. (2016). Novel Mechanisms and Treatment Approaches in Autism Spectrum Disorder. *Discovery Medicine*, 22(119), 47–54.

Brix, M. K., Ersland, L., Hugdahl, K., Grüner, R., Posserud, M.-B., Hammar, Å., Craven, A. R., Noeske, R., Evans, C. J., Walker, H. B., Midtvedt, T., & Beyer, M. K. (2015). “Brain MR spectroscopy in autism spectrum disorder—The GABA excitatory/inhibitory

imbalance theory revisited". *Frontiers in Human Neuroscience*, 9, 365.

<https://doi.org/10.3389/fnhum.2015.00365>

- Brown, M. S., Singel, D., Hepburn, S., & Rojas, D. C. (2013). Increased glutamate concentration in the auditory cortex of persons with autism and first-degree relatives: A 1H-MRS study. *Autism Research : Official Journal of the International Society for Autism Research*, 6(1), 1–10. <https://doi.org/10.1002/aur.1260>
- Caruana, F., Avanzini, P., Gozzo, F., Pelliccia, V., Casaceli, G., & Rizzolatti, G. (2017). A mirror mechanism for smiling in the anterior cingulate cortex. *Emotion (Washington, D.C.)*, 17(2), 187–190. <https://doi.org/10.1037/emo0000237>
- Carvalho Pereira, A., Violante, I. R., Mougá, S., Oliveira, G., & Castelo-Branco, M. (2018). Medial Frontal Lobe Neurochemistry in Autism Spectrum Disorder is Marked by Reduced N-Acetylaspartate and Unchanged Gamma-Aminobutyric Acid and Glutamate + Glutamine Levels. *Journal of Autism and Developmental Disorders*, 48(5), 1467–1482. <https://doi.org/10.1007/s10803-017-3406-8>
- Chauhan, A., Audhya, T., & Chauhan, V. (2012). Brain Region-Specific Glutathione Redox Imbalance in Autism. *Neurochemical Research*, 37(8), 1681–1689. <https://doi.org/10.1007/s11064-012-0775-4>
- Chien, Y.-L., Chen, Y.-C., & Gau, S. S.-F. (2021). Altered cingulate structures and the associations with social awareness deficits and CNTNAP2 gene in autism spectrum disorder. *NeuroImage. Clinical*, 31, 102729. <https://doi.org/10.1016/j.nicl.2021.102729>
- Christensen, D. L., Baio, J., Braun, K. V. N., Bilder, D., Charles, J., Constantino, J. N., Daniels, J., Durkin, M. S., Fitzgerald, R. T., Kurzius-Spencer, M., Lee, L.-C., Pettygrove, S., Robinson, C., Schulz, E., Wells, C., Wingate, M. S., Zahorodny, W., & Yeargin-Allsopp,

M. (2016). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *MMWR Surveillance Summaries*, *65*(3), 1–23.

<https://doi.org/10.15585/mmwr.ss6503a1>

Cochran, D. M., Sikoglu, E. M., Hodge, S. M., Edden, R. A. E., Foley, A., Kennedy, D. N., Moore, C. M., & Frazier, J. A. (2015). Relationship among Glutamine,  $\gamma$ -Aminobutyric Acid, and Social Cognition in Autism Spectrum Disorders. *Journal of Child and Adolescent Psychopharmacology*, *25*(4), 314–322. <https://doi.org/10.1089/cap.2014.0112>

Cogram, P., Deacon, R. M. J., Warner-Schmidt, J. L., von Schimmelmann, M. J., Abrahams, B. S., & During, M. J. (2019). Gaboxadol Normalizes Behavioral Abnormalities in a Mouse Model of Fragile X Syndrome. *Frontiers in Behavioral Neuroscience*, *13*.  
<https://www.frontiersin.org/articles/10.3389/fnbeh.2019.00141>

Cole, E. J., Barraclough, N. E., & Andrews, T. J. (2019). Reduced connectivity between mentalizing and mirror systems in autism spectrum condition. *Neuropsychologia*, *122*, 88–97. <https://doi.org/10.1016/j.neuropsychologia.2018.11.008>

Courchesne, E. (2004). Brain development in autism: Early overgrowth followed by premature arrest of growth. *Mental Retardation and Developmental Disabilities Research Reviews*, *10*(2), 106–111. <https://doi.org/10.1002/mrdd.20020>

Courchesne, E., Karns, C. M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D., Chisum, H. J., Moses, P., Pierce, K., Lord, C., Lincoln, A. J., Pizzo, S., Schreibman, L., Haas, R. H., Akshoomoff, N. A., & Courchesne, R. Y. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology*, *57*(2), 245–254.  
<https://doi.org/10.1212/WNL.57.2.245>

de Giambattista, C., Ventura, P., Trerotoli, P., Margari, F., & Margari, L. (2021). Sex Differences in Autism Spectrum Disorder: Focus on High Functioning Children and Adolescents.

*Frontiers in Psychiatry, 12.*

<https://www.frontiersin.org/articles/10.3389/fpsy.2021.539835>

DeLorey, T. M., Sahbaie, P., Hashemi, E., Homanics, G. E., & Clark, J. D. (2008). Gabrb3 gene deficient mice exhibit impaired social and exploratory behaviors, deficits in non-selective attention and hypoplasia of cerebellar vermal lobules: A potential model of autism spectrum disorder. *Behavioural Brain Research, 187*(2), 207–220.

<https://doi.org/10.1016/j.bbr.2007.09.009>

DeMayo, M. M., Harris, A. D., Song, Y. J. C., Pokorski, I., Thapa, R., Patel, S., Ambarchi, Z., Thomas, E. E., Hickie, I. B., & Guastella, A. J. (2021). Age-related parietal GABA alterations in children with autism spectrum disorder. *Autism Research, 14*(5), 859–872. <https://doi.org/10.1002/aur.2487>

Di Costanzo, A., Trojsi, F., Tosetti, M., Schirmer, T., Lechner, S. M., Popolizio, T., & Scarabino, T. (2007). Proton MR spectroscopy of the brain at 3 T: An update. *European Radiology, 17*(7), 1651–1662. <https://doi.org/10.1007/s00330-006-0546-1>

*Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed* (pp. xlv, 947).

(2013). American Psychiatric Publishing, Inc.

<https://doi.org/10.1176/appi.books.9780890425596>

Dichter, G. S., Felder, J. N., & Bodfish, J. W. (2009). Autism is characterized by dorsal anterior cingulate hyperactivation during social target detection. *Social Cognitive and Affective Neuroscience, 4*(3), 215. <https://doi.org/10.1093/scan/nsp017>



- D’Mello, A. M., Frosch, I. R., Li, C. E., Cardinaux, A. L., & Gabrieli, J. D. E. (2022). Exclusion of females in autism research: Empirical evidence for a “leaky” recruitment-to-research pipeline. *Autism Research, 15*(10), 1929–1940. <https://doi.org/10.1002/aur.2795>
- Donovan, A. P. A., & Basson, M. A. (2017). The neuroanatomy of autism – a developmental perspective. *Journal of Anatomy, 230*(1), 4–15. <https://doi.org/10.1111/joa.12542>
- Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open, 6*(12), e011458. <https://doi.org/10.1136/bmjopen-2016-011458>
- Doyle-Thomas, K. A. R., Card, D., Soorya, L. V., Wang, A. T., Fan, J., & Anagnostou, E. (2014). Metabolic mapping of deep brain structures and associations with symptomatology in autism spectrum disorders. *Research in Autism Spectrum Disorders, 8*(1), 44–51. <https://doi.org/10.1016/j.rasd.2013.10.003>
- Drenthen, G. S., Barendse, E. M., Aldenkamp, A. P., van Veenendaal, T. M., Puts, N. A. J., Edden, R. A. E., Zinger, S., Thoonen, G., Hendriks, M. P. H., Kessels, R. P. C., & Jansen, J. F. A. (2016). Altered neurotransmitter metabolism in adolescents with high-functioning autism. *Psychiatry Research: Neuroimaging, 256*, 44–49. <https://doi.org/10.1016/j.psychresns.2016.09.007>
- Dringen, R. (2000). Metabolism and functions of glutathione in brain. *Progress in Neurobiology, 62*(6), 649–671. [https://doi.org/10.1016/S0301-0082\(99\)00060-X](https://doi.org/10.1016/S0301-0082(99)00060-X)
- Du, F., Zhang, Y., Iltis, I., Marjanska, M., Zhu, X.-H., Henry, P.-G., & Chen, W. (2009). In vivo proton MRS to quantify anesthetic effects of pentobarbital on cerebral metabolism and brain activity in rat. *Magnetic Resonance in Medicine, 62*(6), 1385–1393. <https://doi.org/10.1002/mrm.22146>

Du, Y., Chen, L., Yan, M.-C., Wang, Y.-L., Zhong, X.-L., Xu, C.-X., Li, Y.-B., & Cheng, Y. (2023).

Neurometabolite levels in the brains of patients with autism spectrum disorders: A meta-analysis of proton magnetic resonance spectroscopy studies (N = 1501).

*Molecular Psychiatry*, 1–12. <https://doi.org/10.1038/s41380-023-02079-y>

Durieux, A. M. S., Horder, J., Mendez, M. A., Egerton, A., Williams, S. C. R., Wilson, C. E.,

Spain, D., Murphy, C., Robertson, D., Barker, G. J., Murphy, D. G., & McAlonan, G. M.

(2016). Cortical and subcortical glutathione levels in adults with autism spectrum

disorder. *Autism Research*, 9(4), 429–435. <https://doi.org/10.1002/aur.1522>

Duval, S., & Tweedie, R. (2000). Trim and Fill: A Simple Funnel-Plot–Based Method of Testing

and Adjusting for Publication Bias in Meta-Analysis. *Biometrics*, 56(2), 455–463.

<https://doi.org/10.1111/j.0006-341X.2000.00455.x>

Edden, R. A. E., Crocetti, D., Zhu, H., Gilbert, D. L., & Mostofsky, S. H. (2012). Reduced GABA

Concentration in Attention-Deficit/Hyperactivity Disorder. *Archives of General*

*Psychiatry*, 69(7), 750–753. <https://doi.org/10.1001/archgenpsychiatry.2011.2280>

Edden, R. A. E., Oeltzschner, G., Harris, A. D., Puts, N. A. J., Chan, K., Boer, V. O., Schär, M., &

Barker, P. B. (2016). Prospective frequency correction for macromolecule suppressed

GABA editing experiments at 3T. *Journal of Magnetic Resonance Imaging : JMRI*,

44(6), 1474–1482. <https://doi.org/10.1002/jmri.25304>

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by

a simple, graphical test. *BMJ*, 315(7109), 629–634.

<https://doi.org/10.1136/bmj.315.7109.629>

Endo, T., Shioiri, T., Kitamura, H., Kimura, T., Endo, S., Masuzawa, N., & Someya, T. (2007).

Altered Chemical Metabolites in the Amygdala-Hippocampus Region Contribute to

Autistic Symptoms of Autism Spectrum Disorders. *Biological Psychiatry*, 62(9), 1030–1037. <https://doi.org/10.1016/j.biopsych.2007.05.015>

Endres, D., Tebartz van Elst, L., Meyer, S. A., Feige, B., Nickel, K., Bubl, A., Riedel, A., Ebert, D., Lange, T., Glauche, V., Biscaldi, M., Philipsen, A., Maier, S. J., & Perlov, E. (2017). Glutathione metabolism in the prefrontal brain of adults with high-functioning autism spectrum disorder: An MRS study. *Molecular Autism*, 8(1), 10. <https://doi.org/10.1186/s13229-017-0122-3>

Etherton, M., Földy, C., Sharma, M., Tabuchi, K., Liu, X., Shamloo, M., Malenka, R. C., & Südhof, T. C. (2011). Autism-linked neuroligin-3 R451C mutation differentially alters hippocampal and cortical synaptic function. *Proceedings of the National Academy of Sciences*, 108(33), 13764–13769. <https://doi.org/10.1073/pnas.1111093108>

Fatemi, S. H., Halt, A. R., Strydom, J. M., Kanodia, R., Schulz, S. C., & Realmuto, G. R. (2002). Glutamic acid decarboxylase 65 and 67 kDa proteins are reduced in autistic parietal and cerebellar cortices. *Biological Psychiatry*, 52(8), 805–810. [https://doi.org/10.1016/S0006-3223\(02\)01430-0](https://doi.org/10.1016/S0006-3223(02)01430-0)

Fernandez, A., Meechan, D. W., Karpinski, B. A., Paronett, E. M., Bryan, C. A., Rutz, H. L., Radin, E. A., Lubin, N., Bonner, E. R., Popratiloff, A., Rothblat, L. A., Maynard, T. M., & LaMantia, A.-S. (2019). Mitochondrial Dysfunction Leads to Cortical Under-connectivity and Cognitive Impairment. *Neuron*, 102(6), 1127-1142.e3. <https://doi.org/10.1016/j.neuron.2019.04.013>

Fisher, Z., & Tipton, E. (2015). *robumeta: An R-package for robust variance estimation in meta-analysis* (arXiv:1503.02220). arXiv. <https://doi.org/10.48550/arXiv.1503.02220>

Ford, T. C., & Crewther, D. P. (2016a). A Comprehensive Review of the (1)H-MRS Metabolite Spectrum in Autism Spectrum Disorder. *Frontiers in Molecular Neuroscience*, 9, 14.  
<https://doi.org/10.3389/fnmol.2016.00014>

Ford, T. C., & Crewther, D. P. (2016b). A Comprehensive Review of the 1H-MRS Metabolite Spectrum in Autism Spectrum Disorder. *Frontiers in Molecular Neuroscience*, 9.  
<https://www.frontiersin.org/articles/10.3389/fnmol.2016.00014>

Frahm, J., Merboldt, K.-D., & Hänicke, W. (1987). Localized proton spectroscopy using stimulated echoes. *Journal of Magnetic Resonance (1969)*, 72(3), 502–508.  
[https://doi.org/10.1016/0022-2364\(87\)90154-5](https://doi.org/10.1016/0022-2364(87)90154-5)

Franklin, T. B., Silva, B. A., Perova, Z., Marrone, L., Masferrer, M. E., Zhan, Y., Kaplan, A., Greetham, L., Verrechia, V., Halman, A., Pagella, S., Vyssotski, A. L., Illarionova, A., Grinevich, V., Branco, T., & Gross, C. T. (2017). Prefrontal cortical control of a brainstem social behavior circuit. *Nature Neuroscience*, 20(2), 260–270.  
<https://doi.org/10.1038/nn.4470>

Friedman, N. P., & Robbins, T. W. (2022). The role of prefrontal cortex in cognitive control and executive function. *Neuropsychopharmacology*, 47(1), Article 1.  
<https://doi.org/10.1038/s41386-021-01132-0>

Frye, R. E., DeLaTorre, R., Taylor, H., Slattery, J., Melnyk, S., Chowdhury, N., & James, S. J. (2013). Redox metabolism abnormalities in autistic children associated with mitochondrial disease. *Translational Psychiatry*, 3(6), Article 6.  
<https://doi.org/10.1038/tp.2013.51>

Fujii, E., Mori, K., Miyazaki, M., Hashimoto, T., Harada, M., & Kagami, S. (2010). Function of the frontal lobe in autistic individuals: A proton magnetic resonance spectroscopic

study. *The Journal of Medical Investigation*, 57(1,2), 35–44.

<https://doi.org/10.2152/jmi.57.35>

Fung, L. K., Flores, R. E., Gu, M., Sun, K. L., James, D., Schuck, R. K., Jo, B., Park, J. H., Lee, B.

C., Jung, J. H., Kim, S. E., Saggar, M., Sacchet, M. D., Warnock, G., Khalighi, M. M.,

Spielman, D., Chin, F. T., & Hardan, A. Y. (2021). Thalamic and prefrontal GABA

concentrations but not GABAA receptor densities are altered in high-functioning

adults with autism spectrum disorder. *Molecular Psychiatry*, 26(5), Article 5.

<https://doi.org/10.1038/s41380-020-0756-y>

Gabis, L., Huang, W., Azizian, A., DeVincent, C., Tudorica, A., Kesner-Baruch, Y., Roche, P., &

Pomeroy, J. (2008). 1H-Magnetic Resonance Spectroscopy Markers of Cognitive and

Language Ability in Clinical Subtypes of Autism Spectrum Disorders. *Journal of Child*

*Neurology*, 23(7), 766–774. <https://doi.org/10.1177/0883073808315423>

Gaetz, W., Bloy, L., Wang, D. J., Port, R. G., Blaskey, L., Levy, S. E., & Roberts, T. P. L. (2014).

GABA estimation in the brains of children on the autism spectrum: Measurement

precision and regional cortical variation. *NeuroImage*, 86, 1–9.

<https://doi.org/10.1016/j.neuroimage.2013.05.068>

Ghisleni, C., Bollmann, S., Poil, S.-S., Brandeis, D., Martin, E., Michels, L., O’Gorman, R. L., &

Klaver, P. (2015). Subcortical Glutamate Mediates the Reduction of Short-Range

Functional Connectivity with Age in a Developmental Cohort. *Journal of*

*Neuroscience*, 35(22), 8433–8441. <https://doi.org/10.1523/JNEUROSCI.4375-14.2015>

Gilbert, S. J., Bird, G., Brindley, R., Frith, C. D., & Burgess, P. W. (2008). Atypical recruitment of

medial prefrontal cortex in autism spectrum disorders: An fMRI study of two

executive function tasks. *Neuropsychologia*, 46(9), 2281–2291.

<https://doi.org/10.1016/j.neuropsychologia.2008.03.025>

- Goji, A., Ito, H., Mori, K., Harada, M., Hisaoka, S., Toda, Y., Mori, T., Abe, Y., Miyazaki, M., & Kagami, S. (2017). Assessment of Anterior Cingulate Cortex (ACC) and Left Cerebellar Metabolism in Asperger's Syndrome with Proton Magnetic Resonance Spectroscopy (MRS). *PLOS ONE*, *12*(1), e0169288. <https://doi.org/10.1371/journal.pone.0169288>
- Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised Algorithms for Improved Diagnostic Validity. *Journal of Autism and Developmental Disorders*, *37*(4), 613–627. <https://doi.org/10.1007/s10803-006-0280-1>
- Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O. A., Anney, R., Awashti, S., Belliveau, R., Bettella, F., Buxbaum, J. D., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Christensen, J. H., ... Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*(3), Article 3. <https://doi.org/10.1038/s41588-019-0344-8>
- Gu, F., Chauhan, V., & Chauhan, A. (2013). Impaired synthesis and antioxidant defense of glutathione in the cerebellum of autistic subjects: Alterations in the activities and protein expression of glutathione-related enzymes. *Free Radical Biology and Medicine*, *65*, 488–496. <https://doi.org/10.1016/j.freeradbiomed.2013.07.021>
- Gu, M., Hurd, R., Noeske, R., Baltusis, L., Hancock, R., Sacchet, M. D., Gotlib, I. H., Chin, F. T., & Spielman, D. M. (2018). GABA-Editing with Macromolecule Suppression using an Improved MEGA-SPECIAL Sequence. *Magnetic Resonance in Medicine*, *79*(1), 41–47. <https://doi.org/10.1002/mrm.26691>
- Guo, B., Chen, J., Chen, Q., Ren, K., Feng, D., Mao, H., Yao, H., Yang, J., Liu, H., Liu, Y., Jia, F., Qi, C., Lynn-Jones, T., Hu, H., Fu, Z., Feng, G., Wang, W., & Wu, S. (2019). Anterior

- cingulate cortex dysfunction underlies social deficits in Shank3 mutant mice. *Nature Neuroscience*, 22(8), Article 8. <https://doi.org/10.1038/s41593-019-0445-9>
- Hak, T., van Rhee, H., & Suurmond, R. (2016). *How to Interpret Results of Meta-Analysis* (SSRN Scholarly Paper 3241367). <https://doi.org/10.2139/ssrn.3241367>
- Harada, M., Taki, M. M., Nose, A., Kubo, H., Mori, K., Nishitani, H., & Matsuda, T. (2011). Non-Invasive Evaluation of the GABAergic/Glutamatergic System in Autistic Patients Observed by MEGA-Editing Proton MR Spectroscopy Using a Clinical 3 Tesla Instrument. *Journal of Autism and Developmental Disorders*, 41(4), 447–454. <https://doi.org/10.1007/s10803-010-1065-0>
- Hardan, A. Y., Fung, L. K., Frazier, T., Berquist, S. W., Minshew, N. J., Keshavan, M. S., & Stanley, J. A. (2016). A proton spectroscopy study of white matter in children with autism. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 66, 48–53. <https://doi.org/10.1016/j.pnpbp.2015.11.005>
- Hardan, A. Y., Minshew, N. J., Melhem, N. M., Srihari, S., Jo, B., Bansal, R., Keshavan, M. S., & Stanley, J. A. (2008). An MRI and proton spectroscopy study of the thalamus in children with autism. *Psychiatry Research: Neuroimaging*, 163(2), 97–105. <https://doi.org/10.1016/j.psychresns.2007.12.002>
- Hashemi, E., Ariza, J., Rogers, H., Noctor, S. C., & Martínez-Cerdeño, V. (2017). The Number of Parvalbumin-Expressing Interneurons Is Decreased in the Prefrontal Cortex in Autism. *Cerebral Cortex*, 27(3), 1931–1943. <https://doi.org/10.1093/cercor/bhw021>
- Hashimoto, T., Tayama, M., Miyazaki, M., Yoneda, Y., Yoshimoto, T., Harada, M., Miyoshi, H., Tanouchi, M., & Kuroda, Y. (1997). Differences in Brain Metabolites Between Patients With Autism and Mental Retardation as Detected by in Vivo Localized Proton

Magnetic Resonance Spectroscopy. *Journal of Child Neurology*, 12(2), 91–96.

<https://doi.org/10.1177/088307389701200204>

Hasler, G., van der Veen, J. W., Tumonis, T., Meyers, N., Shen, J., & Drevets, W. C. (2007).

Reduced Prefrontal Glutamate/Glutamine and  $\gamma$ -Aminobutyric Acid Levels in Major Depression Determined Using Proton Magnetic Resonance Spectroscopy. *Archives of General Psychiatry*, 64(2), 193–200. <https://doi.org/10.1001/archpsyc.64.2.193>

Hassan, T. H., Abdelrahman, H. M., Abdel Fattah, N. R., El-Masry, N. M., Hashim, H. M., El-

Gerby, K. M., & Abdel Fattah, N. R. (2013). Blood and brain glutamate levels in children with autistic disorder. *Research in Autism Spectrum Disorders*, 7(4), 541–548. <https://doi.org/10.1016/j.rasd.2012.12.005>

Hau, J., Aljawad, S., Baggett, N., Fishman, I., Carper, R. A., & Müller, R.-A. (2019). The

cingulum and cingulate U-fibers in children and adolescents with autism spectrum disorders. *Human Brain Mapping*, 40(11), 3153–3164.

<https://doi.org/10.1002/hbm.24586>

Havdahl, A., Niarchou, M., Starnawska, A., Uddin, M., van der Merwe, C., & Warrier, V.

(2021). Genetic contributions to autism spectrum disorder. *Psychological Medicine*, 51(13), 2260–2273. <https://doi.org/10.1017/S0033291721000192>

Hazlett, H. C., Poe, M., Gerig, G., Smith, R. G., Provenzale, J., Ross, A., Gilmore, J., & Piven, J.

(2005). Magnetic Resonance Imaging and Head Circumference Study of Brain Size in Autism: Birth Through Age 2 Years. *Archives of General Psychiatry*, 62(12), 1366–1376. <https://doi.org/10.1001/archpsyc.62.12.1366>

Haznedar, M. M., Buchsbaum, M. S., Metzger, M., Solimando, A., Spiegel-Cohen, J., &

Hollander, E. (1997). Anterior cingulate gyrus volume and glucose metabolism in



autistic disorder. *The American Journal of Psychiatry*, 154(8), 1047–1050.

<https://doi.org/10.1176/ajp.154.8.1047>

He, J. L., Oeltzschner, G., Mikkelsen, M., Deronda, A., Harris, A. D., Crocetti, D., Wodka, E. L., Mostofsky, S. H., Edden, R. A. E., & Puts, N. A. J. (2021). Region-specific elevations of glutamate + glutamine correlate with the sensory symptoms of autism spectrum disorders. *Translational Psychiatry*, 11(1), Article 1. <https://doi.org/10.1038/s41398-021-01525-1>

Hegarty, J. P., Gu, M., Spielman, D. M., Cleveland, S. C., Hallmayer, J. F., Lazzeroni, L. C., Raman, M. M., Frazier, T. W., Phillips, J. M., Reiss, A. L., & Hardan, A. Y. (2018). A proton MR spectroscopy study of the thalamus in twins with autism spectrum disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 81, 153–160. <https://doi.org/10.1016/j.pnpbp.2017.09.016>

Higgins, J. P., & Green, S. (2002). *Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series*.

Hisaoka, S., Harada, M., Nishitani, H., & Mori, K. (2001). Regional magnetic resonance spectroscopy of the brain in autistic individuals. *Neuroradiology*, 43(6), 496–498. <https://doi.org/10.1007/s002340000520>

Hollestein, V., Buitelaar, J. K., Brandeis, D., Banaschewski, T., Kaiser, A., Hohmann, S., Oranje, B., Gooskens, B., Durston, S., Williams, S. C. R., Lythgoe, D. J., & Naaijen, J. (2021). Developmental changes in fronto-striatal glutamate and their association with functioning during inhibitory control in autism spectrum disorder and obsessive compulsive disorder. *NeuroImage: Clinical*, 30, 102622. <https://doi.org/10.1016/j.nicl.2021.102622>

- Horder, J., Lavender, T., Mendez, M. A., O’Gorman, R., Daly, E., Craig, M. C., Lythgoe, D. J., Barker, G. J., & Murphy, D. G. (2013). Reduced subcortical glutamate/glutamine in adults with autism spectrum disorders: A [1H]MRS study. *Translational Psychiatry*, 3(7), Article 7. <https://doi.org/10.1038/tp.2013.53>
- Horder, J., Petrinovic, M. M., Mendez, M. A., Bruns, A., Takumi, T., Spooren, W., Barker, G. J., Künnecke, B., & Murphy, D. G. (2018). Glutamate and GABA in autism spectrum disorder—A translational magnetic resonance spectroscopy study in man and rodent models. *Translational Psychiatry*, 8(1), Article 1. <https://doi.org/10.1038/s41398-018-0155-1>
- Ito, H., Mori, K., Harada, M., Hisaoka, S., Toda, Y., Mori, T., Goji, A., Abe, Y., Miyazaki, M., & Kagami, S. (2017). A Proton Magnetic Resonance Spectroscopic Study in Autism Spectrum Disorder Using a 3-Tesla Clinical Magnetic Resonance Imaging (MRI) System: The Anterior Cingulate Cortex and the Left Cerebellum. *Journal of Child Neurology*, 32(8), 731–739. <https://doi.org/10.1177/0883073817702981>
- Jacob, Z., Li, H., Makaryus, R., Zhang, S., Reinsel, R., Lee, H., Feng, T., Rothman, D. L., & Benveniste, H. (2012). Metabolomic Profiling of Children’s Brains Undergoing General Anesthesia with Sevoflurane and Propofol. *Anesthesiology*, 117(5), 1062–1071. <https://doi.org/10.1097/ALN.0b013e31826be417>
- Jiménez-Espinoza, C., Marcano Serrano, F., & González-Mora, J. L. (2021). N-Acetylaspartyl-Glutamate Metabolism in the Cingulated Cortices as a Biomarker of the Etiology in ASD: A 1H-MRS Model. *Molecules*, 26(3), Article 3. <https://doi.org/10.3390/molecules26030675>
- Joshi, G., Biederman, J., Wozniak, J., Goldin, R. L., Crowley, D., Furtak, S., Lukas, S. E., & Gönenç, A. (2013). Magnetic resonance spectroscopy study of the glutamatergic

system in adolescent males with high-functioning autistic disorder: A pilot study at

4T. *European Archives of Psychiatry and Clinical Neuroscience*, 263(5), 379–384.

<https://doi.org/10.1007/s00406-012-0369-9>

Kana, R. K., Keller, T. A., Minshew, N. J., & Just, M. A. (2007). Inhibitory Control in High-Functioning Autism: Decreased Activation and Underconnectivity in Inhibition Networks. *Biological Psychiatry*, 62(3), 198–206.

<https://doi.org/10.1016/j.biopsych.2006.08.004>

Keller, T. A., Kana, R. K., & Just, M. A. (2007). A developmental study of the structural integrity of white matter in autism. *Neuroreport*, 18(1), 23–27.

<https://doi.org/10.1097/01.wnr.0000239965.21685.99>

Kemper, T. L., & Bauman, M. L. (1993). The contribution of neuropathologic studies to the understanding of autism. *Neurologic Clinics*, 11(1), 175–187.

Khacho, M., Clark, A., Svoboda, D. S., MacLaurin, J. G., Lagace, D. C., Park, D. S., & Slack, R. S. (2017). Mitochondrial dysfunction underlies cognitive defects as a result of neural stem cell depletion and impaired neurogenesis. *Human Molecular Genetics*, 26(17), 3327–3341. <https://doi.org/10.1093/hmg/ddx217>

<https://doi.org/10.1093/hmg/ddx217>

Kim, H.-J., Cho, M.-H., Shim, W. H., Kim, J. K., Jeon, E.-Y., Kim, D.-H., & Yoon, S.-Y. (2017). Deficient autophagy in microglia impairs synaptic pruning and causes social behavioral defects. *Molecular Psychiatry*, 22(11), 1576–1584.

<https://doi.org/10.1038/mp.2016.103>

Kirkovski, M., Suo, C., Enticott, P. G., Yücel, M., & Fitzgerald, P. B. (2018). Short communication: Sex-linked differences in gamma-aminobutyric acid (GABA) are related to social functioning in autism spectrum disorder. *Psychiatry Research: Neuroimaging*, 274, 19–22. <https://doi.org/10.1016/j.psychresns.2018.02.004>

<https://doi.org/10.1016/j.psychresns.2018.02.004>

Klein Gunnewiek, T. M., Van Hugte, E. J. H., Frega, M., Guardia, G. S., Foreman, K., Panneman, D., Mossink, B., Linda, K., Keller, J. M., Schubert, D., Cassiman, D., Rodenburg, R., Vidal Folch, N., Oglesbee, D., Perales-Clemente, E., Nelson, T. J., Morava, E., Nadif Kasri, N., & Kozicz, T. (2020). M.3243A > G-Induced Mitochondrial Dysfunction Impairs Human Neuronal Development and Reduces Neuronal Network Activity and Synchronicity. *Cell Reports*, *31*(3), 107538.

<https://doi.org/10.1016/j.celrep.2020.107538>

Kleinhans, N. M., Richards, T., Weaver, K. E., Liang, O., Dawson, G., & Aylward, E. (2009). Brief Report: Biochemical Correlates of Clinical Impairment in High Functioning Autism and Asperger's Disorder. *Journal of Autism and Developmental Disorders*, *39*(7), 1079–1086. <https://doi.org/10.1007/s10803-009-0707-6>

Klintwall, L., Holm, A., Eriksson, M., Carlsson, L. H., Olsson, M. B., Hedvall, Å., Gillberg, C., & Fernell, E. (2011). Sensory abnormalities in autism: A brief report. *Research in Developmental Disabilities*, *32*(2), 795–800.

<https://doi.org/10.1016/j.ridd.2010.10.021>

Knaus, T. A., Tager-Flusberg, H., Mock, J., Dauterive, R., & Foundas, A. L. (2012). Prefrontal and occipital asymmetry and volume in boys with autism spectrum disorder. *Cognitive and Behavioral Neurology: Official Journal of the Society for Behavioral and Cognitive Neurology*, *25*(4), 186–194.

<https://doi.org/10.1097/WNN.0b013e318280e154>

Kolodny, T., Schallmo, M.-P., Gerdtts, J., Edden, R. A. E., Bernier, R. A., & Murray, S. O. (2020). Concentrations of Cortical GABA and Glutamate in Young Adults With Autism Spectrum Disorder. *Autism Research : Official Journal of the International Society for Autism Research*, *13*(7), 1111–1129. <https://doi.org/10.1002/aur.2300>

- Kubas, B., Kuřak, W., Sobaniec, W., Tarasow, E., Ĺebkowska, U., & Walecki, J. (2012).  
Metabolite alterations in autistic children: A 1H MR spectroscopy study. *Advances in Medical Sciences*, 57(1), 152–156. <https://doi.org/10.2478/v10039-012-0014-x>
- Lajeunesse, M. J. (2016). Facilitating systematic reviews, data extraction and meta-analysis with the METAGEAR package for R. *Methods in Ecology and Evolution*, 7(3), 323–330.  
<https://doi.org/10.1111/2041-210X.12472>
- Lemonnier, E., Degrez, C., Phelep, M., Tyzio, R., Josse, F., Grandgeorge, M., Hadjikhani, N., & Ben-Ari, Y. (2012). A randomised controlled trial of bumetanide in the treatment of autism in children. *Translational Psychiatry*, 2(12), e202.  
<https://doi.org/10.1038/tp.2012.124>
- Lemonnier, E., Villeneuve, N., Sonie, S., Serret, S., Rosier, A., Roue, M., Brosset, P., Viellard, M., Bernoux, D., Rondeau, S., Thummler, S., Ravel, D., & Ben-Ari, Y. (2017). Effects of bumetanide on neurobehavioral function in children and adolescents with autism spectrum disorders. *Translational Psychiatry*, 7(3), e1056.  
<https://doi.org/10.1038/tp.2017.10>
- Levitt, J. G., O’Neill, J., Blanton, R. E., Smalley, S., Fadale, D., McCracken, J. T., Guthrie, D., Toga, A. W., & Alger, J. R. (2003). Proton magnetic resonance spectroscopic imaging of the brain in childhood autism. *Biological Psychiatry*, 54(12), 1355–1366.  
[https://doi.org/10.1016/S0006-3223\(03\)00688-7](https://doi.org/10.1016/S0006-3223(03)00688-7)
- Libero, L. E., DeRamus, T. P., Lahti, A. C., Deshpande, G., & Kana, R. K. (2015). Multimodal neuroimaging based classification of autism spectrum disorder using anatomical, neurochemical, and white matter correlates. *Cortex*, 66, 46–59.  
<https://doi.org/10.1016/j.cortex.2015.02.008>

- Libero, L. E., Reid, M. A., White, D. M., Salibi, N., Lahti, A. C., & Kana, R. K. (2016). Biochemistry of the Cingulate Cortex in Autism: An MR Spectroscopy Study. *Autism Research : Official Journal of the International Society for Autism Research*, 9(6), 643–657. <https://doi.org/10.1002/aur.1562>
- Lin, A., Andronesi, O., Bogner, W., Choi, I., Coello, E., Cudalbu, C., Juchem, C., Kemp, G. J., Kreis, R., Krššák, M., Lee, P., Maudsley, A. A., Meyerspeer, M., Mlynarik, V., Near, J., Öz, G., Peek, A. L., Puts, N. A., Ratai, E., ... Mullins, P. G. (2021). Minimum Reporting Standards for in vivo Magnetic Resonance Spectroscopy (MRSinMRS): Experts' consensus recommendations. *Nmr in Biomedicine*, 34(5), e4484. <https://doi.org/10.1002/nbm.4484>
- Lin, L., & Chu, H. (2018). Quantifying Publication Bias in Meta-Analysis. *Biometrics*, 74(3), 785–794. <https://doi.org/10.1111/biom.12817>
- Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, 56(6), 466–474. <https://doi.org/10.1016/j.jaac.2017.03.013>
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*, 30(3), 205–223. <https://doi.org/10.1023/A:1005592401947>
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive

developmental disorders. *Journal of Autism and Developmental Disorders*, 24(5), 659–685. <https://doi.org/10.1007/BF02172145>

Madden, J. M., Lakoma, M. D., Lynch, F. L., Rusinak, D., Owen-Smith, A. A., Coleman, K. J., Quinn, V. P., Yau, V. M., Qian, Y. X., & Croen, L. A. (2017). Psychotropic Medication Use among Insured Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, 47(1), 144–154. <https://doi.org/10.1007/s10803-016-2946-7>

Mahdavinassab, S.-M., Saghadzadeh, A., Motamed-Gorji, N., Vaseghi, S., Mohammadi, M.-R., Alichani, R., & Akhondzadeh, S. (2019). Baclofen as an adjuvant therapy for autism: A randomized, double-blind, placebo-controlled trial. *European Child & Adolescent Psychiatry*, 28(12), 1619–1628. <https://doi.org/10.1007/s00787-019-01333-5>

Maier, S., Düppers, A. L., Runge, K., Dacko, M., Lange, T., Fangmeier, T., Riedel, A., Ebert, D., Endres, D., Domschke, K., Perlov, E., Nickel, K., & Tebartz van Elst, L. (2022). Increased prefrontal GABA concentrations in adults with autism spectrum disorders. *Autism Research*, 15(7), 1222–1236. <https://doi.org/10.1002/aur.2740>

Mapelli, L., Soda, T., D'Angelo, E., & Prestori, F. (2022). The Cerebellar Involvement in Autism Spectrum Disorders: From the Social Brain to Mouse Models. *International Journal of Molecular Sciences*, 23(7), 3894. <https://doi.org/10.3390/ijms23073894>

McAlonan, G. M., Cheung, V., Cheung, C., Suckling, J., Lam, G. Y., Tai, K. S., Yip, L., Murphy, D. G. M., & Chua, S. E. (2005). Mapping the brain in autism. A voxel-based MRI study of volumetric differences and intercorrelations in autism. *Brain*, 128(2), 268–276. <https://doi.org/10.1093/brain/awh332>

Mescher, M., Merkle, H., Kirsch, J., Garwood, M., & Gruetter, R. (1998). Simultaneous in vivo spectral editing and water suppression. *NMR in Biomedicine*, *11*(6), 266–272.

[https://doi.org/10.1002/\(SICI\)1099-1492\(199810\)11:6<266::AID-NBM530>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1099-1492(199810)11:6<266::AID-NBM530>3.0.CO;2-J)

Mescher, M., Tannus, A., Johnson, M. O., & Garwood, M. (1996). Solvent Suppression Using Selective Echo Dephasing. *Journal of Magnetic Resonance, Series A*, *123*(2), 226–229.

<https://doi.org/10.1006/jmra.1996.0242>

Mikkelsen, M., Loo, R. S., Puts, N. A. J., Edden, R. A. E., & Harris, A. D. (2018). Designing GABA-Edited Magnetic Resonance Spectroscopy Studies: Considerations of Scan Duration, Signal-To-Noise Ratio and Sample Size. *Journal of Neuroscience Methods*, *303*, 86–94. <https://doi.org/10.1016/j.jneumeth.2018.02.012>

Miller, B. L., Changl, L., Booth, R., Ernst, T., Cornford, M., Nikas, D., McBride, D., & Jenden, D. J. (1996). In vivo <sup>1</sup>H MRS choline: Correlation with in vitro chemistry/histology. *Life Sciences*, *58*(22), 1929–1935. [https://doi.org/10.1016/0024-3205\(96\)00182-8](https://doi.org/10.1016/0024-3205(96)00182-8)

Mohapatra, A. N., & Wagner, S. (2023). The role of the prefrontal cortex in social interactions of animal models and the implications for autism spectrum disorder. *Frontiers in Psychiatry*, *14*, 1205199. <https://doi.org/10.3389/fpsyt.2023.1205199>

Moreno-De-Luca, D., Sanders, S. J., Willsey, A. J., Mulle, J. G., Lowe, J. K., Geschwind, D. H., State, M. W., Martin, C. L., & Ledbetter, D. H. (2013). Using large clinical data sets to infer pathogenicity for rare copy number variants in autism cohorts. *Molecular Psychiatry*, *18*(10), Article 10. <https://doi.org/10.1038/mp.2012.138>

Mori, K., Toda, Y., Ito, H., Mori, T., Goji, A., Fujii, E., Miyazaki, M., Harada, M., & Kagami, S. (2013). A proton magnetic resonance spectroscopic study in autism spectrum disorders: Amygdala and orbito-frontal cortex. *Brain and Development*, *35*(2), 139–145. <https://doi.org/10.1016/j.braindev.2012.09.016>



- Mostofsky, S. H., Powell, S. K., Simmonds, D. J., Goldberg, M. C., Caffo, B., & Pekar, J. J. (2009). Decreased connectivity and cerebellar activity in autism during motor task performance. *Brain*, *132*(9), 2413–2425. <https://doi.org/10.1093/brain/awp088>
- Moxon-Emre, I., Daskalakis, Z. J., Blumberger, D. M., Croarkin, P. E., Lyon, R. E., Forde, N. J., Tani, H., Truong, P., Lai, M.-C., Desarkar, P., Sailasuta, N., Szatmari, P., & Ameis, S. H. (2021). Modulation of Dorsolateral Prefrontal Cortex Glutamate/Glutamine Levels Following Repetitive Transcranial Magnetic Stimulation in Young Adults With Autism. *Frontiers in Neuroscience*, *15*.  
<https://www.frontiersin.org/articles/10.3389/fnins.2021.711542>
- Mullins, P. G., McGonigle, D. J., O’Gorman, R. L., Puts, N. A. J., Vidyasagar, R., Evans, C. J., & Edden, R. A. E. (2014). Current practice in the use of MEGA-PRESS spectroscopy for the detection of GABA. *NeuroImage*, *86*, 43–52.  
<https://doi.org/10.1016/j.neuroimage.2012.12.004>
- Murphy, D. G. M., Critchley, H. D., Schmitz, N., McAlonan, G., van Amelsvoort, T., Robertson, D., Daly, E., Rowe, A., Russell, A., Simmons, A., Murphy, K. C., & Howlin, P. (2002). Asperger Syndrome: A Proton Magnetic Resonance Spectroscopy Study of Brain. *Archives of General Psychiatry*, *59*(10), 885–891.  
<https://doi.org/10.1001/archpsyc.59.10.885>
- Muthukumaraswamy, S. D., Edden, R. A. E., Jones, D. K., Swettenham, J. B., & Singh, K. D. (2009). Resting GABA concentration predicts peak gamma frequency and fMRI amplitude in response to visual stimulation in humans. *Proceedings of the National Academy of Sciences*, *106*(20), 8356–8361.  
<https://doi.org/10.1073/pnas.0900728106>

- Naaijen, J., Zwiers, M. P., Amiri, H., Williams, S. C. R., Durston, S., Oranje, B., Brandeis, D., Boecker-Schlier, R., Ruf, M., Wolf, I., Banaschewski, T., Glennon, J. C., Franke, B., Buitelaar, J. K., & Lythgoe, D. J. (2017). Fronto-Striatal Glutamate in Autism Spectrum Disorder and Obsessive Compulsive Disorder. *Neuropsychopharmacology*, *42*(12), Article 12. <https://doi.org/10.1038/npp.2016.260>
- Nadeem, A., Ahmad, S. F., Al-Harbi, N. O., Attia, S. M., Alshammari, M. A., Alzahrani, K. S., & Bakheet, S. A. (2019). Increased oxidative stress in the cerebellum and peripheral immune cells leads to exaggerated autism-like repetitive behavior due to deficiency of antioxidant response in BTBR T + tf/J mice. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *89*, 245–253. <https://doi.org/10.1016/j.pnpbp.2018.09.012>
- Nakagawa, S., Lagisz, M., Jennions, M. D., Koricheva, J., Noble, D. W. A., Parker, T. H., Sánchez-Tójar, A., Yang, Y., & O’Dea, R. E. (2022). Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution*, *13*(1), 4–21. <https://doi.org/10.1111/2041-210X.13724>
- O’Brien, F. M., Page, L., O’Gorman, R. L., Bolton, P., Sharma, A., Baird, G., Daly, E., Hallahan, B., Conroy, R. M., Foy, C., Curran, S., Robertson, D., Murphy, K. C., & Murphy, D. G. M. (2010). Maturation of limbic regions in Asperger syndrome: A preliminary study using proton magnetic resonance spectroscopy and structural magnetic resonance imaging. *Psychiatry Research: Neuroimaging*, *184*(2), 77–85. <https://doi.org/10.1016/j.psychresns.2010.08.007>
- O’Neill, J., Bansal, R., Goh, S., Rodie, M., Sawardekar, S., & Peterson, B. S. (2020). Parsing the Heterogeneity of Brain Metabolic Disturbances in Autism Spectrum Disorder. *Biological Psychiatry*, *87*(2), 174–184. <https://doi.org/10.1016/j.biopsych.2019.06.010>

Oner, O., Devrimci-Ozguven, H., Oktem, F., Yagmurlu, B., Baskak, B., & Munir, K. M. (2007).

Proton MR Spectroscopy: Higher Right Anterior Cingulate N-Acetylaspartate/Choline Ratio in Asperger Syndrome Compared with Healthy Controls. *American Journal of Neuroradiology*, 28(8), 1494–1498. <https://doi.org/10.3174/ajnr.A0625>

Otsuka, H., Harada, M., Mori, K., Hisaoka, S., & Nishitani, H. (1999). Brain metabolites in the

hippocampus-amygdala region and cerebellum in autism: An 1H-MR spectroscopy study. *Neuroradiology*, 41(7), 517–519. <https://doi.org/10.1007/s002340050795>

Page, L. A., Daly, E., Schmitz, N., Simmons, A., Toal, F., Deeley, Q., Ambery, F., McAlonan, G.

M., Murphy, K. C., & Murphy, D. G. M. (2006). In Vivo 1 H-Magnetic Resonance Spectroscopy Study of Amygdala-Hippocampal and Parietal Regions in Autism. *American Journal of Psychiatry*, 163(12), 2189–2192.

<https://doi.org/10.1176/ajp.2006.163.12.2189>

Pasanta, D., He, J. L., Ford, T., Oeltzschner, G., Lythgoe, D. J., & Puts, N. A. (2023). Functional

MRS studies of GABA and glutamate/Glx – A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 144, 104940.

<https://doi.org/10.1016/j.neubiorev.2022.104940>

Peek, A. L., Rebbeck, T., Puts, N. A. J., Watson, J., Aguila, M.-E. R., & Leaver, A. M. (2020). Brain

GABA and glutamate levels across pain conditions: A systematic literature review and meta-analysis of 1H-MRS studies using the MRS-Q quality assessment tool.

*NeuroImage*, 210, 116532. <https://doi.org/10.1016/j.neuroimage.2020.116532>

Pierce, S., Kadlaskar, G., Edmondson, D. A., McNally Keehn, R., Dydak, U., & Keehn, B. (2021).

Associations between sensory processing and electrophysiological and neurochemical measures in children with ASD: An EEG-MRS study. *Journal of*

*Neurodevelopmental Disorders*, 13(1), 5. <https://doi.org/10.1186/s11689-020-09351-0>

Pollack, M. H., Jensen, J. E., Simon, N. M., Kaufman, R. E., & Renshaw, P. F. (2008). High-field MRS study of GABA, glutamate and glutamine in social anxiety disorder: Response to treatment with levetiracetam. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 32(3), 739–743. <https://doi.org/10.1016/j.pnpbp.2007.11.023>

Porges, E. C., Jensen, G., Foster, B., Edden, R. A., & Puts, N. A. (2021). The trajectory of cortical GABA across the lifespan, an individual participant data meta-analysis of edited MRS studies. *eLife*, 10, e62575. <https://doi.org/10.7554/eLife.62575>

Port, R. G., Oberman, L. M., & Roberts, T. P. (2019). Revisiting the excitation/inhibition imbalance hypothesis of ASD through a clinical lens. *The British Journal of Radiology*, 92(1101), 20180944. <https://doi.org/10.1259/bjr.20180944>

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prevention Science*, 23(3), 425–438. <https://doi.org/10.1007/s11121-021-01246-3>

Puts, N. A. J., & Edden, R. A. E. (2012). In vivo magnetic resonance spectroscopy of GABA: A methodological review. *Progress in Nuclear Magnetic Resonance Spectroscopy*, 60, 29–41. <https://doi.org/10.1016/j.pnmrs.2011.06.001>

Puts, N. A. J., Wodka, E. L., Harris, A. D., Crocetti, D., Tommerdahl, M., Mostofsky, S. H., & Edden, R. A. E. (2017). Reduced GABA and altered somatosensory function in children with autism spectrum disorder. *Autism Research*, 10(4), 608–619. <https://doi.org/10.1002/aur.1691>

Qi, C., Chen, A., Mao, H., Hu, E., Ge, J., Ma, G., Ren, K., Xue, Q., Wang, W., & Wu, S. (2022). Excitatory and Inhibitory Synaptic Imbalance Caused by Brain-Derived Neurotrophic

Factor Deficits During Development in a Valproic Acid Mouse Model of Autism.

*Frontiers in Molecular Neuroscience*, 15.

<https://www.frontiersin.org/articles/10.3389/fnmol.2022.860275>

Rae, C. D. (2014). A Guide to the Metabolic Pathways and Function of Metabolites Observed in Human Brain 1H Magnetic Resonance Spectra. *Neurochemical Research*, 39(1), 1–36. <https://doi.org/10.1007/s11064-013-1199-5>

Reyngoudt, H., Claeys, T., Vlerick, L., Verleden, S., Acou, M., Deblaere, K., De Deene, Y., Audenaert, K., Goethals, I., & Achten, E. (2012). Age-related differences in metabolites in the posterior cingulate cortex and hippocampus of normal ageing brain: A 1H-MRS study. *European Journal of Radiology*, 81(3), e223–e231. <https://doi.org/10.1016/j.ejrad.2011.01.106>

Robertson, C. E., Ratai, E.-M., & Kanwisher, N. (2016). Reduced GABAergic Action in the Autistic Brain. *Current Biology*, 26(1), 80–85. <https://doi.org/10.1016/j.cub.2015.11.019>

Rogers, S. J., & Ozonoff, S. (2005). Annotation: What do we know about sensory dysfunction in autism? A critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry*, 46(12), 1255–1268. <https://doi.org/10.1111/j.1469-7610.2005.01431.x>

Rohatgi, A. (2022). *WebPlotDigitizer—Extract data from plots, images, and maps*. <https://automeris.io/WebPlotDigitizer/citation.html>

Rojas, D. C., Singel, D., Steinmetz, S., Hepburn, S., & Brown, M. S. (2014). Decreased left perisylvian GABA concentration in children with autism and unaffected siblings. *NeuroImage*, 86, 28–34. <https://doi.org/10.1016/j.neuroimage.2013.01.045>

Rosenthal, R. (1986). *Meta-Analytic Procedures for Social Science Research* Sage

Publications: Beverly Hills, 1984, 148 pp. *Educational Researcher*, 15(8), 18–20.

<https://doi.org/10.3102/0013189X015008018>

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: Increased ratio of

excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, 2(5), 255–267.

<https://doi.org/10.1034/j.1601-183X.2003.00037.x>

Saleh, M. G., Oeltzschner, G., Chan, K. L., Puts, N. A. J., Schär, M., Harris, A. D., & Edden, R. A.

E. (2016). Simultaneous Edited MRS of GABA and Glutathione. *NeuroImage*, 142,

576–582. <https://doi.org/10.1016/j.neuroimage.2016.07.056>

Sampaio-Baptista, C., Filippini, N., Stagg, C. J., Near, J., Scholz, J., & Johansen-Berg, H. (2015).

Changes in functional connectivity and GABA levels with long-term motor learning.

*NeuroImage*, 106, 15–20. <https://doi.org/10.1016/j.neuroimage.2014.11.032>

Sanacora, G., Mason, G. F., Rothman, D. L., & Krystal, J. H. (2002). Increased Occipital Cortex

GABA Concentrations in Depressed Patients After Therapy With Selective Serotonin

Reuptake Inhibitors. *American Journal of Psychiatry*, 159(4), 663–665.

<https://doi.org/10.1176/appi.ajp.159.4.663>

Sanders, S. J. (2015). First glimpses of the neurobiology of autism spectrum disorder. *Current*

*Opinion in Genetics & Development*, 33, 80–92.

<https://doi.org/10.1016/j.gde.2015.10.002>

Sapey-Triomphe, L.-A., Temmerman, J., Puts, N. A. J., & Wagemans, J. (2021). Prediction

learning in adults with autism and its molecular correlates. *Molecular Autism*, 12(1),

64. <https://doi.org/10.1186/s13229-021-00470-6>

- Schmitz, N., Rubia, K., Daly, E., Smith, A., Williams, S., & Murphy, D. G. M. (2006). Neural correlates of executive function in autistic spectrum disorders. *Biological Psychiatry*, *59*(1), 7–16. <https://doi.org/10.1016/j.biopsych.2005.06.007>
- Schmitz, T. W., Correia, M. M., Ferreira, C. S., Prescott, A. P., & Anderson, M. C. (2017). Hippocampal GABA enables inhibitory control over unwanted thoughts. *Nature Communications*, *8*(1), Article 1. <https://doi.org/10.1038/s41467-017-00956-z>
- Siddiqui, M. F., Elwell, C., & Johnson, M. H. (2016). Mitochondrial Dysfunction in Autism Spectrum Disorders. *Autism-Open Access*, *6*(5), 1000190. <https://doi.org/10.4172/2165-7890.1000190>
- Silverman, J. L., Pride, M. C., Hayes, J. E., Puhger, K. R., Butler-Struben, H. M., Baker, S., & Crawley, J. N. (2015). GABAB Receptor Agonist R-Baclofen Reverses Social Deficits and Reduces Repetitive Behavior in Two Mouse Models of Autism. *Neuropsychopharmacology*, *40*(9), 2228–2239. <https://doi.org/10.1038/npp.2015.66>
- Simms, M. L., Kemper, T. L., Timbie, C. M., Bauman, M. L., & Blatt, G. J. (2009). The anterior cingulate cortex in autism: Heterogeneity of qualitative and quantitative cytoarchitectonic features suggests possible subgroups. *Acta Neuropathologica*, *118*(5), 673–684. <https://doi.org/10.1007/s00401-009-0568-2>
- Singhal, N. K., Huang, H., Li, S., Clements, R., Gadd, J., Daniels, A., Kooijman, E. E., Bannerman, P., Burns, T., Guo, F., Pleasure, D., Freeman, E., Shriver, L., & McDonough, J. (2017). The neuronal metabolite NAA regulates histone H3 methylation in oligodendrocytes and myelin lipid composition. *Experimental Brain Research*, *235*(1), 279–292. <https://doi.org/10.1007/s00221-016-4789-z>
- Stagg, C. J., Bachtar, V., Amadi, U., Gudberg, C. A., Ilie, A. S., Sampaio-Baptista, C., O’Shea, J., Woolrich, M., Smith, S. M., Filippini, N., Near, J., & Johansen-Berg, H. (2014). Local

GABA concentration is related to network-level resting functional connectivity. *eLife*, 3, e01465. <https://doi.org/10.7554/eLife.01465>

Sterne, J. A. C., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology*, 54(10), 1046–1055. [https://doi.org/10.1016/S0895-4356\(01\)00377-8](https://doi.org/10.1016/S0895-4356(01)00377-8)

Strang, J. F., Kenworthy, L., Daniolos, P., Case, L., Wills, M. C., Martin, A., & Wallace, G. L. (2012). Depression and Anxiety Symptoms in Children and Adolescents with Autism Spectrum Disorders without Intellectual Disability. *Research in Autism Spectrum Disorders*, 6(1), 406–412. <https://doi.org/10.1016/j.rasd.2011.06.015>

Sumiya, M., Okamoto, Y., Koike, T., Tanigawa, T., Okazawa, H., Kosaka, H., & Sadato, N. (2020). Attenuated activation of the anterior rostral medial prefrontal cortex on self-relevant social reward processing in individuals with autism spectrum disorder. *NeuroImage. Clinical*, 26, 102249. <https://doi.org/10.1016/j.nicl.2020.102249>

Suurmond, R., van Rhee, H., & Hak, T. (2017). Introduction, comparison, and validation of Meta-Essentials: A free and simple tool for meta-analysis. *Research Synthesis Methods*, 8(4), 537–553. <https://doi.org/10.1002/jrsm.1260>

Suzuki, K., Nishimura, K., Sugihara, G., Nakamura, K., Tsuchiya, K. J., Matsumoto, K., Takebayashi, K., Isoda, H., Sakahara, H., Sugiyama, T., Tsujii, M., Takei, N., & Mori, N. (2010). Metabolite alterations in the hippocampus of high-functioning adult subjects with autism. *International Journal of Neuropsychopharmacology*, 13(4), 529–534. <https://doi.org/10.1017/S1461145709990952>

Tang, G., Gudsnuk, K., Kuo, S.-H., Cotrina, M. L., Rosoklija, G., Sosunov, A., Sonders, M. S., Kanter, E., Castagna, C., Yamamoto, A., Yue, Z., Arancio, O., Peterson, B. S., Champagne, F., Dwork, A. J., Goldman, J., & Sulzer, D. (2014). Loss of mTOR-



dependent macroautophagy causes autistic-like synaptic pruning deficits. *Neuron*, 83(5), 1131–1143. <https://doi.org/10.1016/j.neuron.2014.07.040>

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling Complex Meta-analytic Data Structures Using Robust Variance Estimates: A Tutorial in R. *Journal of Developmental and Life-Course Criminology*, 2(1), 85–112. <https://doi.org/10.1007/s40865-016-0026-5>

Tapper, S., Göransson, N., Lundberg, P., Tisell, A., & Zsigmond, P. (2020). A pilot study of essential tremor: Cerebellar GABA+/Glx ratio is correlated with tremor severity. *Cerebellum & Ataxias*, 7(1), 8. <https://doi.org/10.1186/s40673-020-00116-y>

Tebartz van Elst, L., Maier, S., Fangmeier, T., Endres, D., Mueller, G. T., Nickel, K., Ebert, D., Lange, T., Hennig, J., Biscaldi, M., Riedel, A., & Perlov, E. (2014). Disturbed cingulate glutamate metabolism in adults with high-functioning autism spectrum disorder: Evidence in support of the excitatory/inhibitory imbalance hypothesis. *Molecular Psychiatry*, 19(12), Article 12. <https://doi.org/10.1038/mp.2014.62>

Terpstra, M., Henry, P.-G., & Gruetter, R. (2003). Measurement of reduced glutathione (GSH) in human brain using LCModel analysis of difference-edited spectra. *Magnetic Resonance in Medicine*, 50(1), 19–23. <https://doi.org/10.1002/mrm.10499>

Thakkar, K. N., Polli, F. E., Joseph, R. M., Tuch, D. S., Hadjikhani, N., Barton, J. J. S., & Manoach, D. S. (2008). Response monitoring, repetitive behaviour and anterior cingulate abnormalities in autism spectrum disorders (ASD). *Brain*, 131(9), 2464–2478. <https://doi.org/10.1093/brain/awn099>

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods*, 20(3), 375–393.  
<https://doi.org/10.1037/met0000011>

Tuchman, R., & Rapin, I. (2002). Epilepsy in autism. *The Lancet Neurology*, 1(6), 352–358.

[https://doi.org/10.1016/S1474-4422\(02\)00160-6](https://doi.org/10.1016/S1474-4422(02)00160-6)

Turner, K. C., Frost, L., Linsenbardt, D., McIlroy, J. R., & Müller, R.-A. (2006). Atypically diffuse functional connectivity between caudate nuclei and cerebral cortex in autism.

*Behavioral and Brain Functions: BBF*, 2, 34. <https://doi.org/10.1186/1744-9081-2-34>

Umesawa, Y., Atsumi, T., Chakrabarty, M., Fukatsu, R., & Ide, M. (2020). GABA Concentration in the Left Ventral Premotor Cortex Associates With Sensory Hyper-Responsiveness in Autism Spectrum Disorders Without Intellectual Disability. *Frontiers in Neuroscience*, 14. <https://www.frontiersin.org/articles/10.3389/fnins.2020.00482>

van Aert, R. C. M., Wicherts, J. M., & van Assen, M. A. L. M. (2016). Conducting Meta-Analyses Based on p Values. *Perspectives on Psychological Science*, 11(5), 713–729.

<https://doi.org/10.1177/1745691616650874>

van der Graaf, M. (2010). In vivo magnetic resonance spectroscopy: Basic methodology and clinical applications. *European Biophysics Journal*, 39(4), 527–540.

<https://doi.org/10.1007/s00249-009-0517-y>

Vasconcelos, M. M., Brito, A. R., Domingues, R. C., Da Cruz Jr, L. C. H., Gasparetto, E. L., Werner Jr, J., & Gonçalves, J. P. S. (2008). Proton Magnetic Resonance Spectroscopy in School-Aged Autistic Children. *Journal of Neuroimaging*, 18(3), 288–295.

<https://doi.org/10.1111/j.1552-6569.2007.00200.x>

Veenstra-VanderWeele, J., Cook, E. H., King, B. H., Zarevics, P., Cherubini, M., Walton-Bowen, K., Bear, M. F., Wang, P. P., & Carpenter, R. L. (2017). Arbaclofen in Children and Adolescents with Autism Spectrum Disorder: A Randomized, Controlled, Phase 2 Trial. *Neuropsychopharmacology*, 42(7), 1390–1398.

<https://doi.org/10.1038/npp.2016.237>

- Viechtbauer, W. (2005). Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics*, 30(3), 261–293. <https://doi.org/10.3102/10769986030003261>
- Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, 36, 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Wallace, G. L., Dankner, N., Kenworthy, L., Giedd, J. N., & Martin, A. (2010). Age-related temporal and parietal cortical thinning in autism spectrum disorders. *Brain*, 133(12), 3745–3754. <https://doi.org/10.1093/brain/awq279>
- Wallimann, T., Wyss, M., Brdiczka, D., Nicolay, K., & Eppenberger, H. M. (1992). Intracellular compartmentation, structure and function of creatine kinase isoenzymes in tissues with high and fluctuating energy demands: The ‘phosphocreatine circuit’ for cellular energy homeostasis. *Biochemical Journal*, 281(Pt 1), 21–40.
- Wilson, M., Andronesi, O., Barker, P. B., Bartha, R., Bizzi, A., Bolan, P. J., Brindle, K. M., Choi, I.-Y., Cudalbu, C., Dydak, U., Emir, U. E., Gonzalez, R. G., Gruber, S., Gruetter, R., Gupta, R. K., Heerschap, A., Henning, A., Hetherington, H. P., Huppi, P. S., ... Howe, F. A. (2019). Methodological consensus on clinical proton MRS of the brain: Review and recommendations. *Magnetic Resonance in Medicine*, 82(2), 527–550. <https://doi.org/10.1002/mrm.27742>
- Wolfers, T., Floris, D. L., Dinga, R., van Rooij, D., Isakoglou, C., Kia, S. M., Zabihi, M., Llera, A., Chowdanayaka, R., Kumar, V. J., Peng, H., Laidi, C., Batalle, D., Dimitrova, R., Charman, T., Loth, E., Lai, M.-C., Jones, E., Baumeister, S., ... Beckmann, C. F. (2019). From pattern classification to stratification: Towards conceptualizing the heterogeneity of Autism Spectrum Disorder. *Neuroscience & Biobehavioral Reviews*, 104, 240–254. <https://doi.org/10.1016/j.neubiorev.2019.07.010>

- Wood-Downie, H., Wong, B., Kovshoff, H., Cortese, S., & Hadwin, J. A. (2021). Research Review: A systematic review and meta-analysis of sex/gender differences in social interaction and communication in autistic and nonautistic children and adolescents. *Journal of Child Psychology and Psychiatry*, *62*(8), 922–936.  
<https://doi.org/10.1111/jcpp.13337>
- Woodward, N. D., Giraldo-Chica, M., Rogers, B., & Cascio, C. J. (2017). Thalamocortical dysconnectivity in autism spectrum disorder: An analysis of the Autism Brain Imaging Data Exchange. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, *2*(1), 76–84. <https://doi.org/10.1016/j.bpsc.2016.09.002>
- World Health Organization. (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. World Health Organization.  
<https://apps.who.int/iris/handle/10665/37108>
- Yan, J., Oliveira, G., Coutinho, A., Yang, C., Feng, J., Katz, C., Sram, J., Bockholt, A., Jones, I. R., Craddock, N., Cook, E. H., Vicente, A., & Sommer, S. S. (2005). Analysis of the neuroligin 3 and 4 genes in autism and other neuropsychiatric patients. *Molecular Psychiatry*, *10*(4), Article 4. <https://doi.org/10.1038/sj.mp.4001629>
- Yang, X., Si, T., Gong, Q., Qiu, L., Jia, Z., Zhou, M., Zhao, Y., Hu, X., Wu, M., & Zhu, H. (2016). Brain gray matter alterations and associated demographic profiles in adults with autism spectrum disorder: A meta-analysis of voxel-based morphometry studies. *Australian & New Zealand Journal of Psychiatry*, *50*(8), 741–753.  
<https://doi.org/10.1177/0004867415623858>
- Yerys, B. E., Jankowski, K. F., Shook, D., Rosenberger, L. R., Barnes, K. A., Berl, M. M., Ritzl, E. K., VanMeter, J., Vaidya, C. J., & Gaillard, W. D. (2009). The fMRI success rate of children and adolescents: Typical development, epilepsy, attention

deficit/hyperactivity disorder, and autism spectrum disorders. *Human Brain*

*Mapping*, 30(10), 3426–3435. <https://doi.org/10.1002/hbm.20767>

Zablotsky, B. (2017). *Estimated Prevalence of Children With Diagnosed Developmental Disabilities in the United States, 2014–2016*. 291.

Zhang, L., Huang, C.-C., Dai, Y., Luo, Q., Ji, Y., Wang, K., Deng, S., Yu, J., Xu, M., Du, X., Tang, Y., Shen, C., Feng, J., Sahakian, B. J., Lin, C.-P., & Li, F. (2020). Symptom improvement in children with autism spectrum disorder following bumetanide administration is associated with decreased GABA/glutamate ratios. *Translational Psychiatry*, 10, 9. <https://doi.org/10.1038/s41398-020-0692-2>

Zhou, S., & Yu, Y. (2018). Synaptic E-I Balance Underlies Efficient Neural Coding. *Frontiers in Neuroscience*, 12. <https://www.frontiersin.org/articles/10.3389/fnins.2018.00046>

Zhou, Y., Shi, L., Cui, X., Wang, S., & Luo, X. (2016). Functional Connectivity of the Caudal Anterior Cingulate Cortex Is Decreased in Autism. *PLOS ONE*, 11(3), e0151879. <https://doi.org/10.1371/journal.pone.0151879>

Zikopoulos, B., & Barbas, H. (2010). Changes in prefrontal axons may disrupt the network in autism. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 30(44), 14595–14609. <https://doi.org/10.1523/JNEUROSCI.2257-10.2010>