

## TITLE PAGE

Immunological pathways underlying autism: Findings from Mendelian randomization and genetic colocalisation analyses.

### Authors

Christina Dardani<sup>1\*</sup>, Jamie W. Robinson<sup>1,2</sup>, Jie Zheng<sup>1,2</sup>, Aws Sadik<sup>1,3</sup>, Panagiota Pagoni<sup>1,2</sup>, Evie Stergiakouli<sup>1,2</sup>, Renee Gardner<sup>4</sup>, Alexandra Havdahl<sup>5,6,7,8</sup>, Jakob Grove<sup>9,10,11,12</sup>, the iPSYCH Autism Spectrum Disorder working group<sup>9\*</sup>, George Davey Smith<sup>1,2,13</sup>, Sarah Sullivan<sup>1,13</sup>, Beate Leppert<sup>1,2</sup>, Hannah J. Jones<sup>1,2,13</sup>, Stan Zammit<sup>1,13,14</sup>, Golam M. Khandaker<sup>1,2,15,16</sup>, Dheeraj Rai<sup>1,3,13</sup>.

### Author affiliations

<sup>1</sup> Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK.

<sup>2</sup> MRC Integrative Epidemiology Unit, Bristol Medical School, University of Bristol, Bristol, UK

<sup>3</sup> Avon and Wiltshire Partnership NHS Mental Health Trust, Bath, UK.

<sup>4</sup> Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden.

<sup>5</sup> Department of Mental Disorders, Norwegian Institute of Public Health, Oslo, Norway.

<sup>6</sup> Nic Waals Institute, Lovisenberg Diakonale Hospital, Oslo, Norway.

<sup>7</sup> MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK.

<sup>8</sup> PROMENTA Research Center, Department of Psychology, University of Oslo, Oslo, Norway.

<sup>9</sup> The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Aarhus, Denmark.

<sup>10</sup> Department of Biomedicine and Center for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark.

<sup>11</sup> Center for Genomics and Personalized Medicine, CGPM, Aarhus University, Aarhus, Denmark.

<sup>12</sup> Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark.

<sup>13</sup> National Institute for Health Research Bristol Biomedical Research Centre, University Hospitals Bristol and Weston NHS Foundation Trust, University of Bristol, Bristol, UK

<sup>14</sup> Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, UK.

<sup>15</sup> Department of Psychiatry, University of Cambridge, Cambridge, UK.

<sup>16</sup> Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, UK.

\*Corresponding author: Christina Dardani, Population Health Sciences, Bristol Medical School, University of Bristol, Oakfield House, Oakfield Grove, BS8 2BN, Bristol; [christina.dardani@bristol.ac.uk](mailto:christina.dardani@bristol.ac.uk)

- The iPSYCH Autism Spectrum Disorder working group: Anders D. Børglum, Jakob Grove, Thomas Damm Als, Thomas Werge, Preben Bo Mortensen, Marianne Giørtz Pedersen, Carsten Bøcker Pedersen, Ole Mors, Merete Nordentoft, David M. Hougaard, Jonas Bybjerg-Grauholm, Marie Bækvad-Hansen, Christine Søholm Hansen

## ABSTRACT

Emerging evidence implicates the role of inflammation and immunity in autism. However, little is known about the involvement of specific immunological pathways and their causal role. In 18,381 autism cases and 27,969 controls from the PGC and the iPSYCH consortia, we investigated whether 15 cytokines implicated in the differentiation and function of CD4<sup>+</sup> T cell subsets (T<sub>H</sub>1, T<sub>H</sub>2, T<sub>H</sub>9, T<sub>FH</sub>, T<sub>H</sub>17, T<sub>Reg</sub>) could be causally linked to autism. Within a Mendelian randomization framework, we used protein quantitative trait loci (pQTLs; N=1,000-3,394) to assess the effects of genetically proxied levels of plasma cytokines on autism. We additionally used brain cortex expression quantitative trait loci (eQTLs; N=6,601) to investigate whether genetically predicted expression of the genes encoding the cytokines of interest influence autism. We performed colocalisation to assess the possibility that the identified effects were confounded due to Linkage Disequilibrium (LD). We also assessed the possibility of reverse causation. We report consistent evidence for causal effects of genetically predicted levels of IFN- $\gamma$ R1, IL-12R $\beta$ 1 (T<sub>H</sub>1), and IL-4RA, IL-5RA, IL-13RA1 (T<sub>H</sub>2) on autism. We identified brain-specific effects of genetically predicted expression of *IFNGR1*, *IL12RB1*, *IL23A*, which in the case of *IFNGR1* and *IL23A* were additionally supported by evidence suggestive of colocalisation. Findings appeared unlikely to be influenced by reverse causation. Our findings are consistent with a potentially causal effect of T<sub>H</sub>1 and T<sub>H</sub>2 pathway cytokines in autism, and further research is required to elucidate the pathways via which T<sub>H</sub>1 and T<sub>H</sub>2 influence its phenotypic presentation.

**ABSTRACT WORD COUNT:** 242

**MAIN TEXT WORD COUNT:** 3,889

## 1 INTRODUCTION

2 Autism is a common neurodevelopmental condition (~1.5% worldwide prevalence<sup>1</sup>),  
3 influencing multiple areas of functioning across the life course<sup>2</sup>, and is associated with  
4 several comorbidities<sup>3</sup>, poor life outcomes<sup>4</sup> and early mortality<sup>5</sup>. Little is currently known  
5 about the biological pathways contributing to autism and elucidating them may enhance  
6 current understanding on the phenotypic presentation, comorbidities, and life outcomes of  
7 autistic individuals.

8 Emerging evidence implicates inflammation and immunity in autism<sup>6-8</sup>. Nationwide registry-  
9 based studies suggest associations between parental autoimmune conditions, as well as  
10 maternal infections during pregnancy, and offspring autism<sup>9-11</sup>. Moreover, atypical levels of  
11 acute phase proteins (indicators of immune response<sup>12</sup>) in maternal serum during the first  
12 trimester of pregnancy, and in neonatal blood spots have been reported to be associated  
13 with autism<sup>13,14</sup>. Finally, two recent meta-analyses of case-control studies have provided  
14 evidence of atypical concentrations of pro- and anti-inflammatory cytokines in plasma and  
15 serum of autistic individuals versus controls<sup>15,16</sup>.

16 Cytokines are pleiotropic proteins<sup>17</sup>. Their key immune role includes driving differentiation of  
17 naïve CD4<sup>+</sup> T cells into specific subsets, which are characterised by distinct cytokine  
18 products and functions<sup>18,19</sup> (T helper 1 (T<sub>H</sub>1), T helper 2 (T<sub>H</sub>2), T helper 9 (T<sub>H</sub>9), T follicular  
19 helper (T<sub>FH</sub>), T helper 17 (T<sub>H</sub>17) and regulatory T cells (T<sub>Reg</sub>); please see Table 1 for an  
20 overview of inductive and product cytokines as well as subset functions). CD4<sup>+</sup> T cells are  
21 types of T lymphocytes and are orchestrators of anti-viral, autoimmune, and anti-tumour  
22 responses<sup>18-21</sup>.

23 There is increasing evidence suggesting the involvement of CD4<sup>+</sup> T cell subsets and their  
24 signature cytokines in pregnancy outcomes, including recurrent miscarriages and  
25 preeclampsia<sup>22</sup>. In the case of autism, rodent studies suggest an association between  
26 maternal T<sub>H</sub>17 cell populations and their product cytokine IL-17A, and offspring autism-like

27 behaviours<sup>23</sup>. Furthermore, observational studies have found atypical levels of T<sub>H</sub>1 and T<sub>H</sub>2  
28 cytokines in neonatal blood spots<sup>24</sup> as well as brain tissue<sup>25</sup> of autistic individuals, while  
29 there are indications of associations between T<sub>H</sub>17, T<sub>H</sub>2 and T<sub>Reg</sub> cell populations and  
30 gastrointestinal symptoms in autism<sup>26</sup>. Despite evidence suggesting a potential link between  
31 CD4<sup>+</sup> T cells, their signature cytokines, and autism, the question of causality has yet to be  
32 settled, as residual confounding and reverse causation remain key challenges of current  
33 observational evidence.

34 In the present study we investigated the causal influence of genetically proxied cytokines  
35 implicated in the differentiation and function of six major CD4<sup>+</sup> T cell subsets (T<sub>H</sub>1, T<sub>H</sub>2, T<sub>H</sub>9,  
36 T<sub>FH</sub>, T<sub>H</sub>17, T<sub>Reg</sub>) on autism, to elucidate potential distinct immunological mechanisms  
37 underlying the condition. We implemented Mendelian randomization (MR), through an  
38 instrumental variables approach, using single nucleotide polymorphisms (SNPs) associated  
39 with plasma cytokines (protein quantitative trait loci- pQTLs)<sup>27-30</sup> as instruments. The  
40 method, under certain assumptions, intends to estimate causal effects and can overcome  
41 some of the limitations of observational studies, particularly reverse causation and residual  
42 confounding<sup>31</sup>. To gain insights into potential brain-specific effects, we additionally performed  
43 MR using SNPs associated with the expression of genes encoding the cytokines of interest  
44 in the brain cortex (expression quantitative trait loci- eQTLs)<sup>32</sup>. We complemented MR  
45 findings by performing genetic colocalisation analyses to provide evidence for the presence  
46 of shared causal variant(s) influencing levels/expression of the exposure (cytokine) and  
47 autism risk, meaning there might be a shared underlying biological mechanism<sup>33,34</sup>. Finally,  
48 we assessed the possibility of bias due to reverse causation by performing Steiger filtering<sup>35</sup>  
49 and bidirectional MR analyses (genetic liability to autism influencing circulating cytokines)<sup>36</sup>.

50

51

52



## 1 MATERIALS AND METHODS

2 A summary of the analysis plan can be found in Figure 1.

### 3 Data sources and instrument definition

#### 4 *Blood plasma pQTL data*

5 Plasma pQTL data for 15 of the cytokines of interest (Table1) were available in four genome-  
6 wide association studies (GWAS): Sun et al, 2018 (N= 3,301)<sup>27</sup>, Folkersen et al., 2017 (N=  
7 3,394)<sup>29</sup>, Suhre et al., 2017 (N= 1,000)<sup>28</sup>, Emilsson et al., 2018 (N= 5,457)<sup>30</sup>. Supplementary  
8 Table S1 summarises sample sizes, population ancestry and methods for phenotype  
9 definition for each study. Further details on participants, plasma protein measurements, and  
10 genotyping of each study, can be found in the original publications.

11 In previous published work by Zheng et al<sup>37</sup>, genetic instruments across these GWASs were  
12 validated in terms of their consistency (limited agreement of pQTL association estimates  
13 across studies might indicate artefactual associations and therefore violation of the first MR  
14 assumption- Figure 2) and their specificity (pQTLs associated with several proteins can  
15 indicate pleiotropy and therefore potential violation of the third MR assumption- Figure 2).  
16 Further details on the validation protocol can be found in the original publication<sup>37</sup>. We  
17 considered this information important for the appraisal of the MR findings and therefore  
18 extracted 18 pQTLs that were independent ( $r^2 < 0.001$ ; 10,000Kb) and robustly associated  
19 ( $p \leq 5 \times 10^{-08}$ ) with the cytokines of interest and their receptors (Supplementary Table S2 for a  
20 summary of the pQTLs, effect sizes, standard errors, p-values, GWAS source, specificity  
21 and consistency across studies). The only exceptions were instruments for IL-4 Receptor  
22 Subunit Alpha (IL-4RA) and IL-17F which were not included in the Zheng et al.<sup>37</sup> validation  
23 process and therefore were extracted ( $r^2 < 0.001$ ; 10,000kb;  $p \leq 5 \times 10^{-08}$ ) from the Sun et al  
24 GWAS data<sup>27</sup> along with information on their respective regions, necessary for subsequent  
25 colocalisation analyses.

26 All instruments were categorised into cis-acting and trans-acting (Supplementary Table S2).  
27 Instruments were categorised as cis-acting when they were located within proximity ( $\pm 1\text{Mb}$ )  
28 to the cytokine-encoding gene, whereas instruments were categorised as trans-acting when  
29 located outside this window. SNPs acting in cis to the cytokine-encoding gene are more  
30 likely to influence mRNA and protein expression (thus being less pleiotropic)<sup>38</sup>. On the other  
31 hand, trans-acting SNPs, are more likely to be pleiotropic due to their distance from the  
32 cytokine-encoding gene, but their inclusion can potentially increase the proportion of  
33 variance explained in the exposure and the power of the MR analyses<sup>37,38</sup>.

#### 34 ***Brain cortex eQTL data***

35 Brain cortex eQTL data for the genes encoding the cytokines of interest were available in the  
36 largest meta-analysis of brain-derived eQTL datasets (*MetaBrain*), resulting in 6,601 RNA-  
37 seq samples (Supplementary Table S1)<sup>32</sup>. Further details on the study datasets, samples and  
38 genotyping can be found in the original publication.

39 Cis-acting only eQTLs ( $\pm 1\text{Mb}$  within the cytokine encoding gene region) were used for these  
40 analyses. This was because the *MetaBrain* study reported only the statistically significant  
41 trans-eQTLs, without information on the respective regions around them. This means that  
42 any genes with trans-acting SNPs are ineligible for subsequent colocalisation analyses. We  
43 defined as instruments SNPs that were independent ( $r^2 < 0.001$ ; 10,000 kb) and met a p-value  
44 threshold of  $5 \times 10^{-08}$ . In cases that there were no instruments available for a cytokine of  
45 interest at this threshold, we used a relaxed p-value threshold of  $5 \times 10^{-05}$  in order to ensure  
46 that there would be at least one cis-acting eQTL for all the cytokine-encoding genes. In total,  
47 19 eQTLs were extracted and details can be found in Supplementary Table S3.

#### 48 ***Autism GWAS data***

49 We used summary-level data from the latest autism GWAS of 18,381 cases and 27,969  
50 controls<sup>39</sup> (Supplementary Table S1). Considering emerging evidence suggesting that  
51 autism with co-occurring intellectual disabilities is distinct to autism without (in terms of



52 behavioural characteristics<sup>40</sup>, genetic and environmental risk factors<sup>41,42</sup>, and comorbid  
53 medical and mental health conditions<sup>43,44</sup>), we additionally used summary-level data on a  
54 sub-sample of the iPSYCH cohort<sup>45</sup> excluding all intellectual disability cases (cases= 11,203;  
55 controls= 22,555).

## 56 **Methods**

### 57 ***Two-sample Mendelian randomization***

58 MR utilises the special properties of germline genetic variants to strengthen causal inference  
59 within observational data<sup>46</sup>. Here we implemented this as an instrumental variables analysis  
60 using SNPs as instruments. The method can yield unbiased causal effect estimates under  
61 assumptions that the instruments should satisfy: (1) they must be robustly associated with  
62 the exposure, (2) they must not be associated with any confounders of the exposure-  
63 outcome associations, (3) they should operate on the outcome entirely through the exposure  
64 (i.e., no horizontal pleiotropy)<sup>47</sup> (Figure 2).

65 For the present study, we performed two-sample MR, in which instrument-exposure and  
66 instrument-outcome effect sizes and standard errors were extracted from separate GWASs  
67 conducted in independent samples but representative of the same underlying population<sup>48</sup>.

68 We assessed the strength of each instrument by estimating their F-statistic, where an F  
69 statistic of >10 is indicative of adequate instrument strength<sup>49</sup>. SNP-exposure effect sizes  
70 and standard errors were extracted from the autism GWAS<sup>39</sup>, and their alleles were  
71 harmonised to ensure SNP-exposure and SNP-outcome effect sizes correspond to the same  
72 allele. The Wald ratio was used to generate causal effect estimates and the two term Taylor  
73 expansion to approximate standard errors, as all exposures were instrumented by a single  
74 SNP<sup>32,50</sup>. The same process was followed using as an outcome the iPSYCH autism sub-  
75 sample excluding all intellectual disability cases. Supplementary Tables S4, S5, S6, S7  
76 contain information on the harmonised datasets used for each MR analysis. We did not  
77 apply correction for multiple testing because the cytokines assessed are organised into

78 interacting subsets (Table 1), limiting therefore the definition of the number of independent  
79 tests. Instead, we orient the readers to appraise the study results in the context of their  
80 consistency across analyses<sup>51</sup>.

### 81 ***Genetic colocalisation***

82 Colocalisation approaches can complement MR approaches by elucidating a distinct aspect  
83 of the identified causal relationship between an exposure and an outcome<sup>34</sup>. Specifically,  
84 colocalisation allows the assessment of the hypothesis that any identified causal effects are  
85 driven by the same causal variant influencing both exposure and outcome, instead of distinct  
86 causal variants that are in linkage disequilibrium (LD) with each other<sup>33</sup>. In practice, the  
87 approach is harnessing SNP coverage within the same specified locus for two traits of  
88 interest and tests whether the association signals for each trait at the specified locus are  
89 suggestive of a shared causal variant<sup>33</sup>.

90 For each MR result providing evidence of a causal effect, we tested for colocalisation  
91 between the genetically proxied exposure and autism. We extracted regions of SNPs within  
92  $\pm 500\text{KB}$  around the instrumented SNP and implemented the algorithm described by Zheng  
93 et al<sup>37</sup> to perform pairwise conditional and colocalisation (PWCoCo) analyses, which  
94 assesses all conditionally independent signals in the exposure dataset region against all  
95 conditionally independent signals in the outcome data. Genotype data from mothers in the  
96 Avon Longitudinal Study of Parents and Children (ALSPAC)<sup>52,53</sup> cohort were used as the LD  
97 reference panel (N= 7,921; details on the ALSPAC cohort and available genotype data can  
98 be found in Supplementary Note S1). We ran these analyses using the default settings, as  
99 suggested by the authors in the original publications<sup>33,37</sup>. Evidence of colocalisation was  
100 considered an H4 (probability of both traits having a shared causal variant)  $\geq 0.8$ , as has  
101 been suggested by the original authors. However, we observed post-hoc that many results  
102 showed high evidence of H0-H2, which assumes no evidence of a causal variant in either  
103 trait possibly due to underpowered datasets. On this basis, we considered as evidence  
104 suggestive of colocalisation a weighted  $H4 \geq 0.8^{54}$ , calculated as  $H4/(H3+H4)$  (the probability

105 of having a shared causal variant for both traits weighted to the probability of having non-  
106 shared causal variants for each trait). The weighted H4 may provide evidence of  
107 colocalization in underpowered datasets but is reliant on the assumption that such a variant  
108 exists.

109 As a post-hoc analysis to PWCoCo and the weighted H4 approach, we also performed LD  
110 check<sup>37</sup>. Although the method was initially proposed to approximate colocalisation in cases  
111 that sufficient SNP coverage in the region was not available<sup>37</sup>, in the context of the present  
112 study it was implemented because the current autism GWAS yielded association signals in  
113 only three loci<sup>39</sup> (assuming that these three loci are unlikely to be the only causal loci for  
114 autism and that future, larger, GWASs will reveal more information on the genetic  
115 architecture of autism). We assessed the LD between the instrumented SNP and the top 30  
116 SNPs associated with autism in the test region ( $r^2 > 0.8$  with any of the strongest 30 SNPs for  
117 autism in the region approximating colocalisation).

#### 118 ***Examination of reverse causation: Steiger filtering***

119 We performed Steiger filtering to assess whether causal effect estimates were influenced by  
120 reverse causation<sup>35</sup>. The method assesses whether the genetic variants proxying for the  
121 exposure explain more variance in the outcome, which, if this is the case, suggests that the  
122 primary phenotype influenced by the variant is the outcome rather than the exposure.

#### 123 ***Examination of reverse causation: Bi-directional two-sample MR***

124 We assessed the causal effects of genetic liability to autism on levels of plasma cytokines.  
125 Ten independent ( $r^2 < 0.01$ ; 10,000 kb) SNPs for autism were extracted from the autism  
126 GWAS<sup>39</sup> using a relaxed p-value threshold of  $\leq 5 \times 10^{-07}$  to increase the number of instruments  
127 included in the analyses ( $p \leq 5 \times 10^{-08}$ ,  $r^2 < 0.01$ ; 10,000 kb, yielded only 2 SNPs)  
128 (Supplementary Table S8 for details on the effect sizes, standard errors and p-values of the  
129 autism instruments). They were then extracted from the Sun et al<sup>27</sup> GWAS summary data for  
130 each cytokine of interest and their alleles were harmonised (Supplementary Table S9 for

131 harmonised datasets). The primary method of analysis was the Inverse Variance Weighted  
132 (IVW)<sup>55</sup>. The consistency of the IVW effect estimates was assessed using the MR Egger  
133 regression<sup>55</sup>, the Weighted Median<sup>56</sup> and the Weighted Mode<sup>57</sup> methods. Details on each  
134 method can be found in Supplementary Note S2. Bi-directional MR analyses could not be  
135 performed using the *MetaBrain* dataset due to not having full genome-wide data available.

## 136 **Software**

137 Analyses were carried out using the computational facilities of the Advanced Computing  
138 Research Centre of the University of Bristol (<http://www.bris.ac.uk/acrc/>). Brain cortex cis-  
139 eQTLs were extracted using the Summary-data-based Mendelian Randomization (SMR)  
140 package version 1.03 (<https://cnsgenomics.com/software/smr/>). The TwoSampleMR R  
141 package was used to conduct two-sample MR analyses, Steiger filtering and to construct LD  
142 matrices for LD check analyses (<https://github.com/MRCIEU/TwoSampleMR>). The PWCoCo  
143 algorithm was implemented using the Pair-Wise Conditional analysis and Colocalisation  
144 analysis package version 0.3 (<https://github.com/jwr-git/pwcoco>).

## 1 RESULTS

### 2 Two-sample Mendelian randomization

#### 3 ***Causal effects of genetically proxied plasma cytokines on autism***

4 All instruments proxying plasma cytokines had good strength ( $F > 10$ , Supplementary Table  
5 S2).

6 We found evidence of a causal effect of genetically proxied Interferon Gamma Receptor-1  
7 (IFN- $\gamma$ R1: OR= 1.15; 95%CI: 1.03-1.29;  $p = 0.02$ ), Interleukin 13 Receptor Subunit Alpha-1  
8 (IL-13RA1: OR= 1.16; 95%CI: 1.00-1.34;  $p = 0.04$ ), Interleukin 4 Receptor Subunit Alpha  
9 (IL4-RA: OR= 0.81; 95%CI: 0.65-0.99;  $p = 0.04$ ) and Interleukin 5 Receptor Subunit Alpha  
10 (IL-5RA: OR= 0.91; 95%CI: 0.83-1.00;  $p = 0.05$ ). There was sub-threshold evidence to  
11 suggest a causal effect of Interleukin 2 (IL2: OR= 1.14; 95%CI: 0.99-1.32;  $p = 0.07$ ) and  
12 Interleukin 12 Receptor Subunit Beta-1 (IL-12R $\beta$ 1: OR= 1.03; 95%CI: 0.99-1.07;  $p = 0.12$ ).  
13 Analyses using the autism sub-sample excluding all intellectual disability cases, yielded  
14 comparable effect estimates. Results are summarised in Supplementary Tables S10 & S11  
15 and visualised in Figure 3.

#### 16 ***Causal effects of brain-expressed cytokine-encoding genes on autism***

17 All instruments had good strength ( $F > 10$ , Supplementary Table S3).

18 There was evidence to suggest a causal effect of genetically predicted expression of  
19 *IFNGR1* gene in brain cortex (OR= 1.22; 95%CI: 1.05- 1.42;  $p = 0.008$ ), and *IL23A* gene  
20 (OR= 0.88; 95%CI: 0.77-0.99;  $p = 0.04$ ). There was suggestive evidence for a potential  
21 causal effect of genetically predicted expression of *IL12B* (OR= 1.24; 95%CI: 0.97-1.57;  $p =$   
22 0.08) and *IL12RB1* gene (OR= 1.10; 95%CI: 0.98-1.23;  $p = 1.11$ ), which was stronger in  
23 subsequent analyses using the autism sub-sample excluding intellectual disability cases  
24 (*IL12B*: OR= 1.36; 95%CI: 1.01- 1.83;  $p = 0.04$ ; *IL12RB1*: OR= 1.16; 95%CI: 1.01-1.34;  $p =$   
25 0.04). Results are summarised in Supplementary Tables S12 & S13 and visualised in Figure  
26 4.

## 27 **Genetic colocalisation**

28 None of the identified effects were supported by evidence of colocalisation i.e.,  $H4 \geq 0.8$ .  
29 However, there was evidence suggestive of colocalisation for the identified causal effect of  
30 genetically predicted expression of *IFNGR1* and *IL23A* in the brain cortex on autism.  
31 Specifically, the weighted  $H4$  ( $H4/(H3+H4)$ ) for both *IFNGR1* and autism as well as *IL23A*  
32 and autism was 0.9 and 0.8 respectively. In addition, LD Check analyses indicated that the  
33 lead *IFNGR1* and *IL23A* variants were in strong LD with at least one of the autism lead  
34 variants in the respective regions ( $r^2 > 0.8$ ) (Supplementary Table S14).

## 35 **Analyses to assess bias due to reverse causation**

36 Steiger filtering indicated that across all analyses the genetic variants explained more  
37 variance in the exposure rather than the outcome, and that therefore the MR causal effect  
38 estimates were unlikely to be influenced by reverse causation (Supplementary Tables S10-  
39 S13).

40 We did not find any causal effect of genetic liability to autism on plasma cytokine levels  
41 (Supplementary Table S15).

## 1 DISCUSSION

### 2 Summary of findings

3 In the present study, we used MR and genetic colocalisation approaches to investigate the  
4 causal influence of genetically proxied cytokines implicated in the differentiation and function  
5 of six major CD4<sup>+</sup> subsets (T<sub>H</sub>1, T<sub>H</sub>2, T<sub>H</sub>9, T<sub>FH</sub>, T<sub>H</sub>17, T<sub>Reg</sub>) on autism, and to elucidate  
6 potentially distinct immunological mechanisms underlying the condition. We found evidence  
7 suggesting a causal effect of genetically proxied T<sub>H</sub>1 (IFN- $\gamma$ R1, IL-12R $\beta$ 1), and T<sub>H</sub>2 (IL-4RA,  
8 IL-5RA, IL13-RA1) signature cytokines. There was additional evidence to suggest a causal  
9 effect of genetically predicted expression of *IFNGR1*, *IL12RB1* and *IL23A* genes in the brain  
10 cortex and especially in the case of *IFNGR1* and *IL23A*, the findings were supported by  
11 evidence suggestive of colocalisation. The identified effects appeared unlikely to be  
12 influenced by reverse causation bias.

### 13 Findings in the context of existing evidence

14 There is a substantial body of evidence suggesting a potentially central role of T<sub>H</sub>1 and T<sub>H</sub>2  
15 signature cytokines in autism. Specifically, in 1,100 neonatal dried blood spots from the  
16 Danish Newborn Screening Biobank, atypical levels of T<sub>H</sub>1 and T<sub>H</sub>2 cytokines (including IFN-  
17  $\gamma$ , IL-2, IL-12, IL-4, IL-5) were found to be associated with later autism diagnosis<sup>24</sup>, while in  
18 1,029 amniotic fluid samples from a Danish historic birth cohort, atypical levels of IL4 and IL5  
19 were found to be associated with autism and broadly childhood neurodevelopmental and  
20 psychiatric disorders<sup>58</sup>. Although these findings could have been potentially influenced by  
21 selection bias (samples in both studies were restricted to pregnancies that amniocentesis  
22 was performed, i.e., high risk pregnancies), recent evidence from the population-based Early  
23 Markers for Autism (EMA) study further supports the potential role of T<sub>H</sub>1 and T<sub>H</sub>2 cytokines  
24 in autism as elevated concentrations of IFN- $\gamma$ , IL-4 and IL-5 in maternal serum during  
25 gestation were found to be associated with offspring autism and intellectual disabilities<sup>59,60</sup>. A  
26 similar pattern has been identified in peripheral blood of children with autism, characterised

27 by atypical levels of IFN- $\gamma$ , IL2, IL4, IL5 and IL13<sup>61,62</sup>, as well as post-mortem brain tissue of  
28 adults with autism, characterised by increased levels of IFN- $\gamma$  and an atypical IFN- $\gamma$ /IL4  
29 ratio<sup>25</sup>. By implementing the principles of MR and overcoming some of the limitations of  
30 previous studies (particularly reverse causation and residual confounding<sup>31</sup>), our study  
31 provides further support for a potentially causal role of genetically proxied T<sub>H</sub>1 and T<sub>H</sub>2  
32 signature cytokines on autism. Especially in the case of IL-12R $\beta$ 1 and IL-5RA, we used cis-  
33 acting SNPs that were specific to the proteins, thus minimising the possibility of bias from  
34 pleiotropy. Similarly in the case of IFN- $\gamma$ R1, the instrument used showed specificity to the  
35 protein and consistency across pQTL studies.

### 36 **Potential immunological pathways in autism**

37 Across MR analyses, a consistent pattern was identified for IFN- $\gamma$ R1. Elevated levels of  
38 genetically proxied IFN- $\gamma$ R1 as well as increased genetically predicted expression of  
39 *IFNGR1* gene in the brain cortex, were found to have causal links with autism. These  
40 findings were further supported by evidence suggestive of colocalisation, as well as  
41 evidence of causal effects of genetically predicted *IL12RB1* and *IL23A* in the brain cortex.

42 Specifically, IL-12R $\beta$ 1 (IL-12/23p40 subunit) is a common receptor for IL-12 and IL-23, which  
43 is promoting their signalling pathways<sup>63</sup>. However, the effect of IL-12R $\beta$ 1 on IL-12 and IL-23  
44 signalling is not uniform. Increasing evidence suggests that IL12RB1 drives naïve CD4<sup>+</sup> T  
45 cell differentiation to T<sub>H</sub>1 (IL12 pathway) or instead to T<sub>H</sub>17 (IL23 pathway), depending on  
46 the presence or absence of Interferon Regulatory Factor 1 (IRF1)<sup>64</sup>. IL12RB1, in the  
47 presence of IRF1, drives naïve CD 4<sup>+</sup> T cell differentiation towards T<sub>H</sub>1 pathway leading to  
48 production of IFN- $\gamma$ , whereas in the absence of IRF1, it drives differentiation towards T<sub>H</sub>17<sup>65</sup>.  
49 Our findings are consistent with this immunobiological understanding, as we identified  
50 evidence that autism is associated with increased genetically predicted expression of  
51 *IL12RB1* and *IFNGR1* in the brain cortex, but with decreased genetically predicted  
52 expression of *IL23A*.



53 Interestingly, the effects of *IFNGR1* and *IL23A* were pronounced in the autism sample  
54 including intellectual disability cases. IFN- $\gamma$  signalling has been found to have a central role  
55 in brain function, influencing neurogenesis, synaptic plasticity and neurodegeneration<sup>66</sup>.  
56 Animal studies indicate that excess IFN- $\gamma$  signalling and production drives neuronal cell  
57 death and synapse loss<sup>67</sup>, while epidemiological studies suggest associations between high  
58 circulating levels of IFN-  $\gamma$  and white matter damage in preterm infants<sup>68</sup>. This evidence  
59 might support our finding of a more pronounced effect of *IFNGR1* expression in brain cortex  
60 in autism cases including intellectual disabilities. However, given the sample sizes of the two  
61 autism GWASs we used, the possibility that our results reflect differences in power cannot  
62 be excluded.

### 63 **Strengths and limitations**

64 The present study benefited from utilising a systematic approach for the selection of immune  
65 markers (based on CD4<sup>+</sup> T cell subsets), as well as from the use of cis-acting genetic  
66 variants proxying for gene expression in the brain cortex. This allowed us to appraise our  
67 findings in the context of underlying immunological pathways and their mechanisms of  
68 action. Furthermore, we implemented a combination of MR and colocalisation approaches to  
69 strengthen causal inference and performed a series of sensitivity analyses to assess the  
70 possibility of reverse causation.

71 However, our findings should be appraised in the context of their limitations. First and  
72 foremost, none of the identified MR effects were supported by robust evidence of  
73 colocalisation. This might suggest that our MR findings were confounded due to LD.

74 Although there was some evidence suggestive of colocalisation based on PWCoCo and  
75 post-hoc LD check analyses, this evidence relied on the assumption that there are causal  
76 variants in the regions of interest which is difficult to ascertain given the possibility of the  
77 datasets being underpowered. Future larger GWASs are necessary to further elucidate the  
78 present colocalisation findings. Second, some of the instruments used in our analyses were  
79 trans-acting, not specific to the cytokines of interest or consistent across studies and were

80 selected using a relaxed p-value threshold. The inclusion of pleiotropic and weak  
81 instruments might have introduced bias in the causal effect estimates<sup>69,70</sup>. Third, although  
82 Steiger filtering suggested that our analyses were unlikely to be influenced by reverse  
83 causation, bi-directional MR analyses may have been underpowered considering the sample  
84 size of the outcome GWAS. Fourth, we assessed the contribution of common genetic  
85 variation and not rare, for which there is evidence of enrichment in immune-function gene  
86 sets in autism<sup>71</sup>. Fifth, we did not have access to family and individual level data which could  
87 have allowed the assessment of the origins of the identified effects (parental vs individual) as  
88 well as the possibility of non-linear effects (which can be particularly relevant in the case of  
89 immune response<sup>72,73</sup>). Sixth, autism is a highly heterogeneous condition and the possibility  
90 of distinct immunological pathways causally influencing autism subtypes was not possible to  
91 be assessed and cannot be excluded. Finally, analyses were conducted using summary  
92 data of European ancestry individuals, limiting therefore the generalisability of the present  
93 findings and replication across ancestries is necessary e.g., Zheng et al.,2021<sup>74</sup>.

## 94 **Conclusions**

95 In conclusion, we found evidence consistent with a causal effect of genetically proxied T<sub>H</sub>1  
96 and T<sub>H</sub>2 signature cytokines on autism. Particularly for IFN- $\gamma$ R1, there was additional MR  
97 and colocalisation evidence to suggest brain-specific effects of its respective gene  
98 expression on autism. The present findings appear unlikely to be influenced by reverse  
99 causation. Further research is necessary in order to elucidate the origins of the identified  
100 effects, the possibility of non-linear relationships and their possible distinct contributions  
101 across autism subtypes.

## 1     **DATA AVAILABILITY STATEMENT**

2     Summary data used to conduct the present analyses can be found at:

3     <https://gwas.mrcieu.ac.uk/>, <https://metabrain.nl>, [https://www.med.unc.edu/pgc/download-](https://www.med.unc.edu/pgc/download-results/)  
4     [results/](https://www.med.unc.edu/pgc/download-results/). Summary data on autism excluding intellectual disabilities can be obtained after  
5     application to the iPSYCH Autism Spectrum Disorder working group.

## 6     **ETHICS DECLARATIONS**

7     Across all analyses, summary-level data were used. Details on ethical approval and  
8     participant informed consent can be found in the original publications<sup>27–30,32,39</sup>. JWR receives  
9     funding from Biogen for research unrelated to the present study. No funding body has  
10    influenced data collection, analyses or their interpretation.

## 11    **ACKNOWLEDGEMENTS**

12    The Medical Research Council (MRC) and the University of Bristol support the MRC  
13    Integrative Epidemiology Unit [, MC\_UU\_00011/1, MC\_UU\_00011/3, MC\_UU\_00011/5]. The  
14    UK Medical Research Council and Wellcome (Grant ref: 217065/Z/19/Z) and the University  
15    of Bristol provide core support for ALSPAC. GWAS data was generated by Sample Logistics  
16    and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory  
17    Corporation of America) using support from 23andMe. A comprehensive list of grants  
18    funding is available on the ALSPAC website  
19    (<http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf>). This  
20    research was funded in part, by the Wellcome Trust. For the purpose of Open Access, the  
21    author has applied a CC BY public copyright licence to any Author Accepted Manuscript  
22    version arising from this submission. CD acknowledges the support of Wellcome Trust  
23    [215379/Z/19/Z]. GDS, HJ, DR, SS, SZ are supported by the NIHR Biomedical Research  
24    Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University  
25    of Bristol. The views expressed are those of the authors and not necessarily those of the  
26    NIHR or the Department of Health and Social Care. GMK acknowledges funding support

27 from the Wellcome Trust (201486/Z/16/Z), the MQ: Transforming Mental Health (grant code:  
28 MQDS17/40), the Medical Research Council UK (grant code: MC\_PC\_17213 and grant  
29 code: MR/S037675/1), NIHR (project code: NIHR202646), and the BMA Foundation (J  
30 Moulton grant 2019). JWR is funded by Biogen. The iPSYCH team was supported by grants  
31 from the Lundbeck Foundation (R102-A9118, R155-2014-1724, and R248-2017-2003),  
32 NIMH (1U01MH109514-01) and the Universities and University Hospitals of Aarhus and  
33 Copenhagen. The Danish National Biobank resource was supported by the Novo Nordisk  
34 Foundation. High-performance computer capacity for handling and statistical analysis of  
35 iPSYCH data on the GenomeDK HPC facility was provided by the Center for Genomics and  
36 Personalized Medicine and the Centre for Integrative Sequencing, iSEQ, Aarhus University,  
37 Denmark. RG acknowledges funding support from the Swedish Research Council (VR2017-  
38 02900). AH was supported by grants from the South-Eastern Norway Regional Health  
39 Authority (2020022, 2018059) and the Research Council of Norway (274611, 288083).

## 1 REFERENCES

- 2 1. Lyall, K. *et al.* The changing epidemiology of autism spectrum disorders. *Annu. Rev. Public Health* **38**, 81–102 (2017).
- 3
- 4 2. Association, A. P. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. (American Psychiatric Pub, 2013).
- 5
- 6 3. Bauman, M. L. Medical comorbidities in autism: challenges to diagnosis and treatment. *Neurotherapeutics* **7**, 320–327 (2010).
- 7
- 8 4. Graham Holmes, L. *et al.* A Lifespan Approach to Patient-Reported Outcomes and Quality of Life for People on the Autism Spectrum. *Autism Res.* **13**, 970–987 (2020).
- 9
- 10 5. Hirvikoski, T. *et al.* Premature mortality in autism spectrum disorder. *Br. J. Psychiatry* **208**, 232–238 (2016).
- 11
- 12 6. Meyer, U., Feldon, J. & Dammann, O. Schizophrenia and autism: both shared and disorder-specific pathogenesis via perinatal inflammation? *Pediatr. Res.* **69**, 26–33 (2011).
- 13
- 14
- 15 7. Hafizi, S., Tabatabaei, D. & Lai, M.-C. Review of clinical studies targeting inflammatory pathways for individuals with autism. *Front. psychiatry* **10**, 849 (2019).
- 16
- 17 8. Thom, R. P. *et al.* Beyond the brain: a multi-system inflammatory subtype of autism spectrum disorder. *Psychopharmacology (Berl)*. **236**, 3045–3061 (2019).
- 18
- 19 9. Lee, B. K. *et al.* Maternal hospitalization with infection during pregnancy and risk of autism spectrum disorders. *Brain. Behav. Immun.* **44**, 100–105 (2015).
- 20
- 21 10. Sadik, A. *et al.* Parental inflammatory bowel disease and autism in the offspring: Triangulating the evidence using four complementary study designs. *medRxiv* 2021.06.09.21258393 (2021) doi:10.1101/2021.06.09.21258393.
- 22
- 23
- 24 11. Rom, A. L. *et al.* Parental rheumatoid arthritis and autism spectrum disorders in offspring: a Danish nationwide cohort study. *J. Am. Acad. Child Adolesc. Psychiatry* **57**, 28–32 (2018).
- 25
- 26
- 27 12. Gabay, C. & Kushner, I. Acute-phase proteins and other systemic responses to inflammation. *N. Engl. J. Med.* **340**, 448–454 (1999).
- 28
- 29 13. Gardner, R. M. *et al.* Neonatal levels of acute phase proteins and risk of autism Spectrum disorder. *Biol. Psychiatry* **89**, 463–475 (2021).
- 30
- 31 14. Brynge, M., Gardner, R. M., Sjöqvist, H., Karlsson, H. & Dalman, C. Maternal Levels of Acute Phase Proteins in Early Pregnancy and Risk of Autism Spectrum Disorders in Offspring. *medRxiv* (2021).
- 32
- 33
- 34 15. Saghazadeh, A. *et al.* A meta-analysis of pro-inflammatory cytokines in autism spectrum disorders: Effects of age, gender, and latitude. *J. Psychiatr. Res.* **115**, 90–102 (2019).
- 35
- 36
- 37 16. Saghazadeh, A. *et al.* Anti-inflammatory cytokines in autism spectrum disorders: A systematic review and meta-analysis. *Cytokine* **123**, 154740 (2019).
- 38
- 39 17. Chauhan, P. *et al.* A primer on cytokines. *Cytokine* **145**, 155458 (2021).
- 40 18. Swain, S. L., McKinstry, K. K. & Strutt, T. M. Expanding roles for CD4+ T cells in immunity to viruses. *Nat. Rev. Immunol.* **12**, 136–148 (2012).
- 41
- 42 19. DuPage, M. & Bluestone, J. A. Harnessing the plasticity of CD4+ T cells to treat

- 43 immune-mediated disease. *Nat. Rev. Immunol.* **16**, 149–163 (2016).
- 44 20. Dittel, B. N. CD4 T cells: balancing the coming and going of autoimmune-mediated  
45 inflammation in the CNS. *Brain. Behav. Immun.* **22**, 421–430 (2008).
- 46 21. Tay, R. E., Richardson, E. K. & Toh, H. C. Revisiting the role of CD4+ T cells in  
47 cancer immunotherapy—new insights into old paradigms. *Cancer Gene Ther.* **28**, 5–  
48 17 (2021).
- 49 22. Wang, W., Sung, N., Gilman-Sachs, A. & Kwak-Kim, J. T helper (Th) cell profiles in  
50 pregnancy and recurrent pregnancy losses: Th1/Th2/Th9/Th17/Th22/Tfh cells. *Front.*  
51 *Immunol.* **11**, 2025 (2020).
- 52 23. Choi, G. B. *et al.* The maternal interleukin-17a pathway in mice promotes autism-like  
53 phenotypes in offspring. *Science (80-. ).* **351**, 933–939 (2016).
- 54 24. Abdallah, M. W. *et al.* Neonatal levels of cytokines and risk of autism spectrum  
55 disorders: an exploratory register-based historic birth cohort study utilizing the Danish  
56 Newborn Screening Biobank. *J. Neuroimmunol.* **252**, 75–82 (2012).
- 57 25. Li, X. *et al.* Elevated immune response in the brain of autistic patients. *J.*  
58 *Neuroimmunol.* **207**, 111–116 (2009).
- 59 26. Rose, D. R. *et al.* T cell populations in children with autism spectrum disorder and co-  
60 morbid gastrointestinal symptoms. *Brain, Behav. Immunity-Health* **2**, 100042 (2020).
- 61 27. Sun, B. B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79  
62 (2018).
- 63 28. Suhre, K. *et al.* Connecting genetic risk to disease end points through the human  
64 blood plasma proteome. *Nat. Commun.* **8**, 1–14 (2017).
- 65 29. Folkersen, L. *et al.* Mapping of 79 loci for 83 plasma protein biomarkers in  
66 cardiovascular disease. *PLoS Genet.* **13**, e1006706 (2017).
- 67 30. Emilsson, V. *et al.* Co-regulatory networks of human serum proteins link genetics to  
68 disease. *Science (80-. ).* **361**, 769–773 (2018).
- 69 31. Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal  
70 inference in epidemiological studies. *Hum. Mol. Genet.* **23**, R89–R98 (2014).
- 71 32. de Klein, N. *et al.* Brain expression quantitative trait locus and network analysis  
72 reveals downstream effects and putative drivers for brain-related diseases. *bioRxiv*  
73 (2021).
- 74 33. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic  
75 association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
- 76 34. Burgess, S., Foley, C. N. & Zuber, V. Inferring causal relationships between risk  
77 factors and outcomes from genome-wide association study data. *Annu. Rev.*  
78 *Genomics Hum. Genet.* **19**, 303–327 (2018).
- 79 35. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between  
80 imprecisely measured traits using GWAS summary data. *PLoS Genet.* **13**, e1007081  
81 (2017).
- 82 36. Holmes, M. V & Davey Smith, G. Can Mendelian randomization shift into reverse  
83 gear? *Clinical chemistry* vol. 65 363–366 (2019).
- 84 37. Zheng, J. *et al.* Phenome-wide Mendelian randomization mapping the influence of the  
85 plasma proteome on complex diseases. *Nat. Genet.* **52**, 1122–1131 (2020).

- 86 38. Swerdlow, D. I. *et al.* Selecting instruments for Mendelian randomization in the wake  
87 of genome-wide association studies. *Int. J. Epidemiol.* **45**, 1600–1616 (2016).
- 88 39. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum  
89 disorder. *Nat. Genet.* **1** (2019).
- 90 40. Kurzius-Spencer, M. *et al.* Behavioral problems in children with autism spectrum  
91 disorder with and without co-occurring intellectual disability. *Res. Autism Spectr.*  
92 *Disord.* **56**, 61–71 (2018).
- 93 41. Xie, S. *et al.* The Familial Risk of Autism Spectrum Disorder with and without  
94 Intellectual Disability. *Autism Res.* (2020).
- 95 42. Rai, D. *et al.* Parental depression, maternal antidepressant use during pregnancy, and  
96 risk of autism spectrum disorders: population based case-control study. *Bmj* **346**,  
97 f2059 (2013).
- 98 43. Amiet, C. *et al.* Epilepsy in autism is associated with intellectual disability and gender:  
99 evidence from a meta-analysis. *Biol. Psychiatry* **64**, 577–582 (2008).
- 100 44. Rai, D. *et al.* Association between autism spectrum disorders with or without  
101 intellectual disability and depression in young adulthood. *JAMA Netw. open* **1**,  
102 e181465–e181465 (2018).
- 103 45. Pedersen, C. B. *et al.* The iPSYCH2012 case-cohort sample: new directions for  
104 unravelling genetic and environmental architectures of severe mental disorders. *Mol.*  
105 *Psychiatry* **23**, 6–14 (2018).
- 106 46. Davey Smith, G. & Ebrahim, S. ‘Mendelian randomization’: can genetic epidemiology  
107 contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.*  
108 **32**, 1–22 (2003).
- 109 47. Haycock, P. C. *et al.* Best (but oft-forgotten) practices: the design, analysis, and  
110 interpretation of Mendelian randomization studies. *Am. J. Clin. Nutr.* **103**, 965–978  
111 (2016).
- 112 48. Burgess, S. *et al.* Using published data in Mendelian randomization: a blueprint for  
113 efficient identification of causal risk factors. *Eur. J. Epidemiol.* **30**, 543–552 (2015).
- 114 49. Pierce, B. L. & Burgess, S. Efficient design for Mendelian randomization studies:  
115 subsample and 2-sample instrumental variable estimators. *Am. J. Epidemiol.* **178**,  
116 1177–1184 (2013).
- 117 50. Wald, A. The fitting of straight lines if both variables are subject to error. *Ann. Math.*  
118 *Stat.* **11**, 284–300 (1940).
- 119 51. Sterne, J. A. C. & Davey Smith, G. Sifting the evidence—what’s wrong with  
120 significance tests? *BMJ* **322**, 226 (2001).
- 121 52. Fraser, A. *et al.* Cohort profile: the Avon Longitudinal Study of Parents and Children:  
122 ALSPAC mothers cohort. *Int. J. Epidemiol.* **42**, 97–110 (2013).
- 123 53. Boyd, A. *et al.* Cohort profile: the ‘children of the 90s’—the index offspring of the Avon  
124 Longitudinal Study of Parents and Children. *Int. J. Epidemiol.* **42**, 111–127 (2013).
- 125 54. Guo, H. *et al.* Integration of disease association and eQTL data using a Bayesian  
126 colocalisation approach highlights six candidate causal genes in immune-mediated  
127 diseases. *Hum. Mol. Genet.* **24**, 3305–3313 (2015).
- 128 55. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid  
129 instruments: effect estimation and bias detection through Egger regression. *Int. J.*



- 130            *Epidemiol.* **44**, 512–525 (2015).
- 131    56.    Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent estimation in  
132            Mendelian randomization with some invalid instruments using a weighted median  
133            estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
- 134    57.    Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data  
135            Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.*  
136            **46**, 1985–1998 (2017).
- 137    58.    Abdallah, M. W. *et al.* Amniotic fluid inflammatory cytokines: potential markers of  
138            immunologic dysfunction in autism spectrum disorders. *World J. Biol. Psychiatry* **14**,  
139            528–538 (2013).
- 140    59.    Goines, P. E. *et al.* Increased midgestational IFN- $\gamma$ , IL-4 and IL-5 in women bearing a  
141            child with autism: a case-control study. *Mol. Autism* **2**, 1–11 (2011).
- 142    60.    Jones, K. L. *et al.* Autism with intellectual disability is associated with increased levels  
143            of maternal cytokines and chemokines during gestation. *Mol. Psychiatry* **22**, 273–279  
144            (2017).
- 145    61.    Gupta, S., Aggarwal, S., Roshanravan, B. & Lee, T. Th1-and Th2-like cytokines in  
146            CD4+ and CD8+ T cells in autism. *J. Neuroimmunol.* **85**, 106–109 (1998).
- 147    62.    Molloy, C. A. *et al.* Elevated cytokine levels in children with autism spectrum disorder.  
148            *J. Neuroimmunol.* **172**, 198–205 (2006).
- 149    63.    Teng, M. W. L. *et al.* IL-12 and IL-23 cytokines: from discovery to targeted therapies  
150            for immune-mediated inflammatory diseases. *Nat. Med.* **21**, 719–729 (2015).
- 151    64.    Kano, S. *et al.* The contribution of transcription factor IRF1 to the interferon- $\gamma$ -  
152            interleukin 12 signaling axis and TH 1 versus TH-17 differentiation of CD4+ T cells.  
153            *Nat. Immunol.* **9**, 34–41 (2008).
- 154    65.    Robinson, R. T. IL12R $\beta$ 1: The cytokine receptor that we used to know. *Cytokine* **71**,  
155            348–359 (2015).
- 156    66.    Monteiro, S., Roque, S., Marques, F., Correia-Neves, M. & Cerqueira, J. J. Brain  
157            interference: revisiting the role of IFN $\gamma$  in the central nervous system. *Prog. Neurobiol.*  
158            **156**, 149–163 (2017).
- 159    67.    Ivashkiv, L. B. IFN $\gamma$ : signalling, epigenetics and roles in immunity, metabolism,  
160            disease and cancer immunotherapy. *Nat. Rev. Immunol.* **18**, 545–558 (2018).
- 161    68.    Hansen-Pupp, I. *et al.* Circulating interferon-gamma and white matter brain damage in  
162            preterm infants. *Pediatr. Res.* **58**, 946–952 (2005).
- 163    69.    Lawlor, D. A. Commentary: Two-sample Mendelian randomization: opportunities and  
164            challenges. *Int. J. Epidemiol.* **45**, 908 (2016).
- 165    70.    Davey Smith, G. & Hemani, G. Mendelian randomization: genetic anchors for causal  
166            inference in epidemiological studies. *Hum. Mol. Genet.* **23**, R89–R98 (2014).
- 167    71.    Ashitha, S. N. M. & Ramachandra, N. B. Integrated functional analysis implicates  
168            syndromic and rare copy number variation genes as prominent molecular players in  
169            pathogenesis of autism spectrum disorders. *Neuroscience* **438**, 25–40 (2020).
- 170    72.    Mayer, H., Zaenker, K. S. & An Der Heiden, U. A basic mathematical model of the  
171            immune response. *Chaos An Interdiscip. J. Nonlinear Sci.* **5**, 155–161 (1995).
- 172    73.    Gutnikov, S. & Melnikov, Y. A simple non-linear model of immune response. *Chaos*,



- 173            *Solitons & Fractals* **16**, 125–132 (2003).
- 174    74.    Zheng, J. *et al.* Trans-ethnic Mendelian-randomization study reveals causal  
175            relationships between cardiometabolic factors and chronic kidney disease. *Int. J.*  
176            *Epidemiol.* **50**, 1995–2010 (2021).
- 177    75.    Angkasekwinai, P. & Dong, C. IL-9-producing T cells: Potential players in allergy and  
178            cancer. *Nat. Rev. Immunol.* **21**, 37–48 (2021).
- 179    76.    Crotty, S. T follicular helper cell differentiation, function, and roles in disease.  
180            *Immunity* **41**, 529–542 (2014).
- 181    77.    Bettelli, E., Korn, T., Oukka, M. & Kuchroo, V. K. Induction and effector functions of  
182            TH 17 cells. *Nature* **453**, 1051–1057 (2008).
- 183    78.    Lucca, L. E. & Dominguez-Villar, M. Modulation of regulatory T cell function and  
184            stability by co-inhibitory receptors. *Nat. Rev. Immunol.* **20**, 680–693 (2020).
- 185

## FIGURE LEGENDS

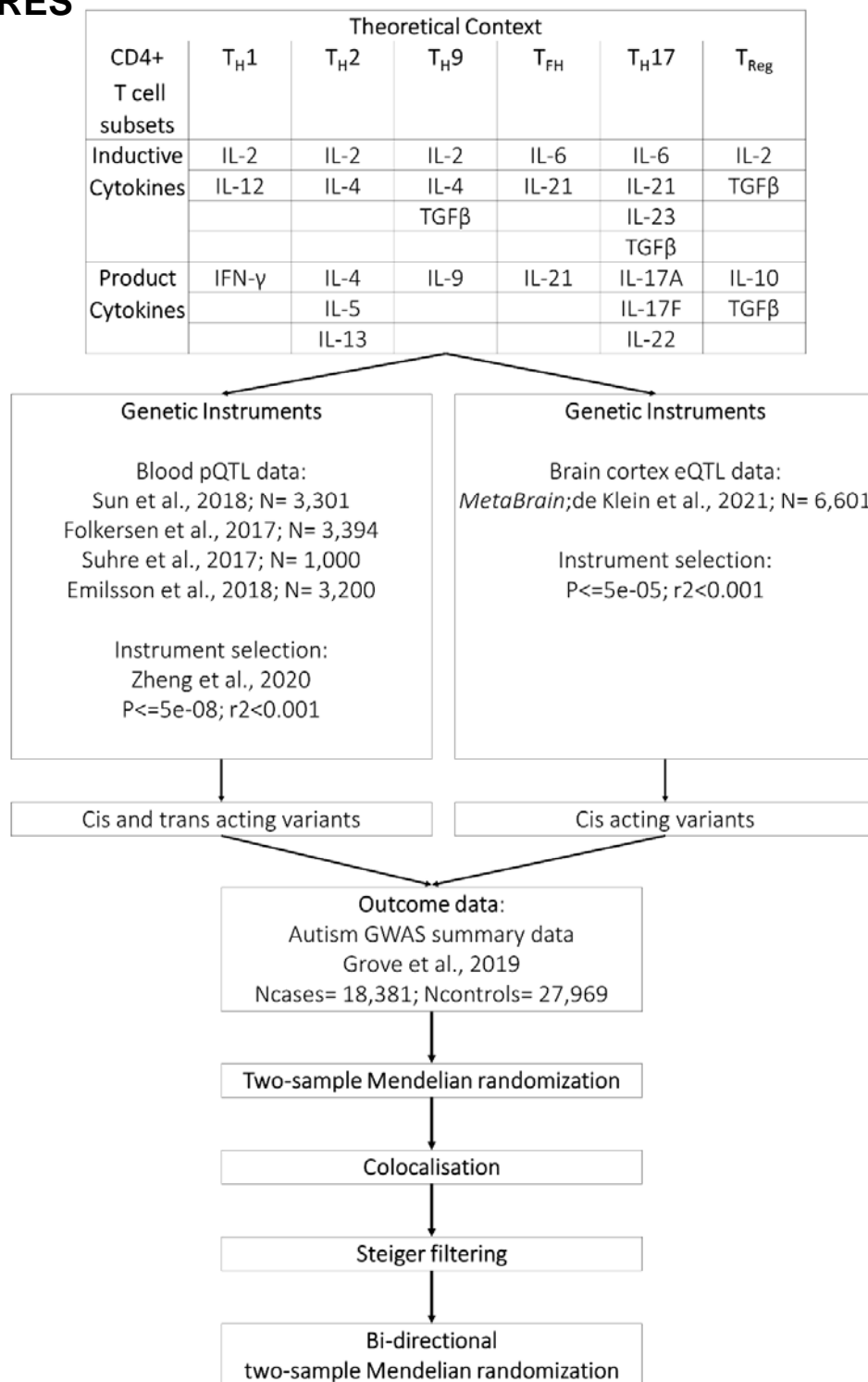
**Figure1.** Summary of the analysis plan followed in the present study.

**Figure 2.** Directed acyclic graph (DAG) visualising the three MR assumptions. Specifically, the method can yield unbiased causal effect estimates under assumptions that the instruments should satisfy: they must be robustly associated with the exposure (MR assumption 1), they must not be associated with any confounders of the exposure-outcome associations (MR assumption 2), they should operate on the outcome entirely through the exposure (i.e. no horizontal pleiotropy).

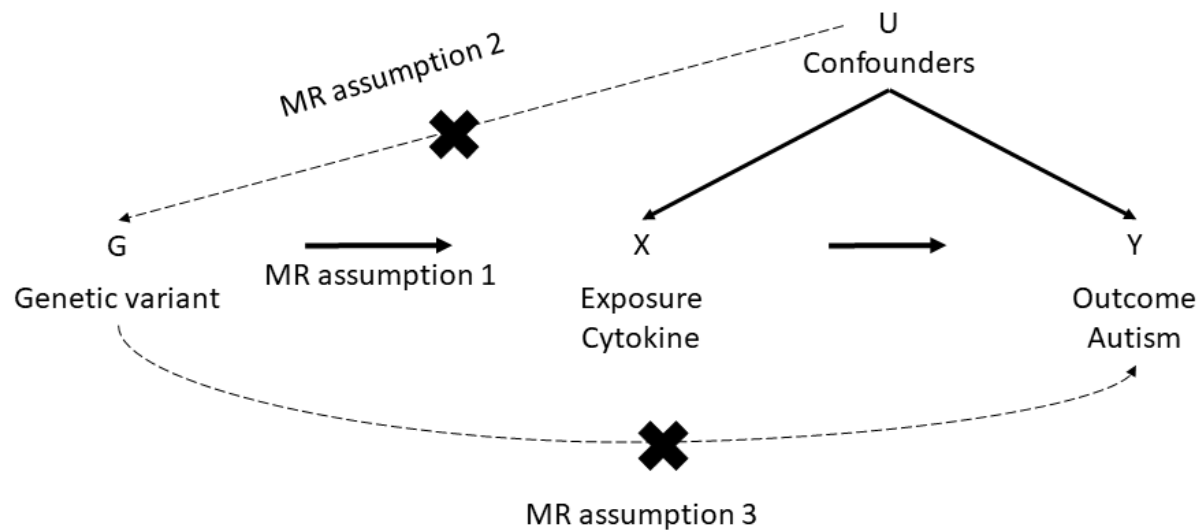
**Figure3.** Forest plot of MR causal effect estimates and 95% CIs for autism per unit change in plasma cytokine levels. IFNGR1: Interferon Gamma Receptor-1; IL10RB: Interleukin 10 Receptor Subunit Beta; IL12B: Interleukin 12 Beta; IL12RB1: Interleukin 12 Receptor Subunit Beta-1; IL12RB2: Interleukin 12 Receptor Subunit Beta-2; IL13RA1: Interleukin 13 Receptor Subunit Alpha-1; IL17F: Interleukin 17 F; IL17RA: Interleukin 17 Receptor Alpha; IL21: Interleukin 21; IL22RA1: Interleukin 22 Receptor Subunit Alpha-1; IL23R: Interleukin 23 Receptor; IL2: Interleukin 2; IL4RA: Interleukin 4 receptor Subunit Alpha; IL5: Interleukin 5; IL6R: Interleukin 6 Receptor; TGFB1: Transforming Growth Factor Beta-1.

**Figure4.** Forest plot of MR causal effect estimates and 95% CIs for autism per standard deviation change in cytokine-encoding gene expression in the brain cortex.

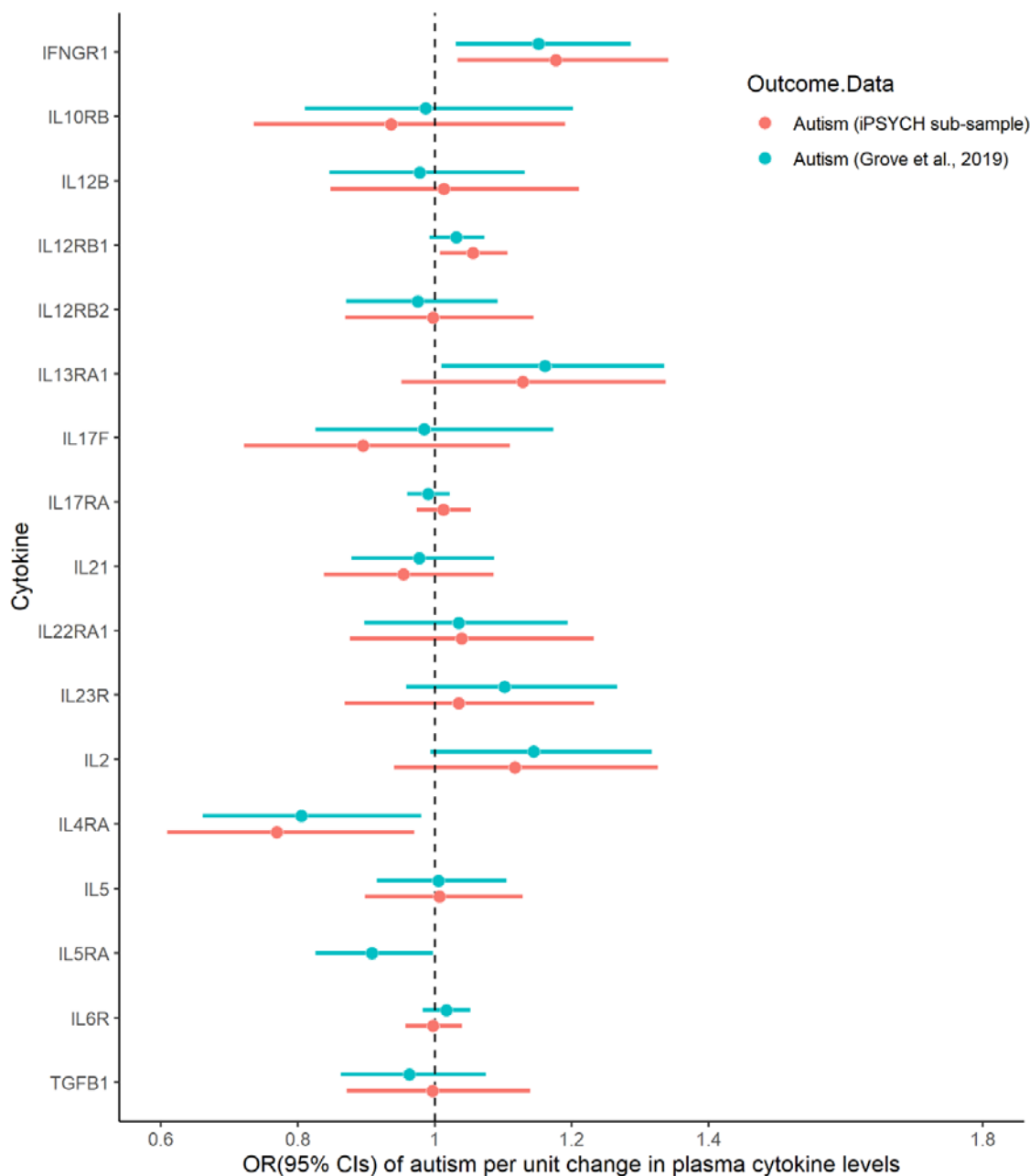
## FIGURES



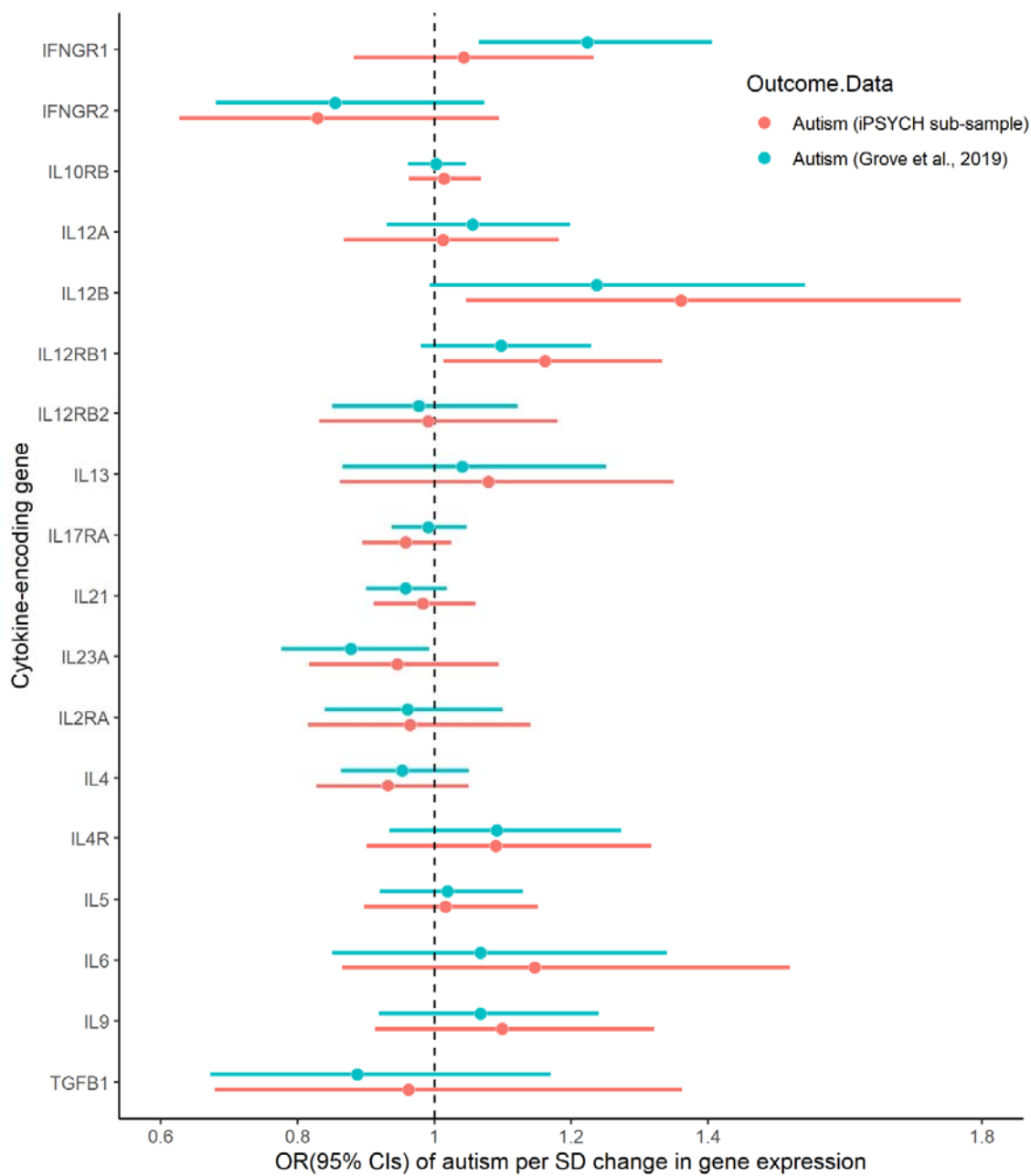
**Figure1.** Summary of the analysis plan followed in the present study.



**Figure 2.** Directed acyclic graph (DAG) visualising the three MR assumptions. Specifically, the method can yield unbiased causal effect estimates under assumptions that the instruments should satisfy: they must be robustly associated with the exposure (MR assumption 1), they must not be associated with any confounders of the exposure-outcome associations (MR assumption 2), they should operate on the outcome entirely through the exposure (i.e. no horizontal pleiotropy).



**Figure3.** Forest plot of MR causal effect estimates and 95% CIs for autism per unit change in plasma cytokine levels. IFNGR1: Interferon Gamma Receptor-1; IL10RB: Interleukin 10 Receptor Subunit Beta; IL12B: Interleukin 12 Beta; IL12RB1: Interleukin 12 Receptor Subunit Beta-1; IL12RB2: Interleukin 12 Receptor Subunit Beta-2; IL13RA1: Interleukin 13 Receptor Subunit Alpha-1; IL17F: Interleukin 17 F; IL17RA: Interleukin 17 Receptor Alpha; IL21: Interleukin 21; IL22RA1: Interleukin 22 Receptor Subunit Alpha-1; IL23R: Interleukin 23 Receptor; IL2: Interleukin 2; IL4RA: Interleukin 4 receptor Subunit Alpha; IL5: Interleukin 5; IL6R: Interleukin 6 Receptor; TGFB1: Transforming Growth Factor Beta-1



**Figure4.** Forest plot of MR causal effect estimates and 95% CIs for autism per standard deviation change in cytokine-encoding gene expression in the brain cortex.

## TABLES

**Table 1.** Summary of the six major CD4<sup>+</sup> T cells, the cytokines inducing their differentiation, their product cytokines and each subset functions\*.

CD4 <sup>+</sup> T cell subsets	T <sub>H</sub> 1	T <sub>H</sub> 2	T <sub>H</sub> 9	T <sub>FH</sub>	T <sub>H</sub> 17	T <sub>Reg</sub>
Inductive Cytokines	IL-2 IL-12	IL-2 IL-4	IL-2 IL-4 TGFβ	IL-6 IL-21	IL-6 IL-21 IL-23 TGFβ	IL-2 TGFβ
Product Cytokines	IFN-γ	IL-4 IL-5 IL-13	IL-9	IL-21	IL-17A IL-17F IL-22	IL-10 TGFβ
Subset functions*	Macrophage activation  Inflammatory response against intracellular pathogens	Eosinophil activation  Allergic and autoimmune response	Response in helminth infections  Allergic and autoimmune response  Anti-tumour immune response	B cell activation  Inflammatory response against extracellular pathogens	Neutrophil activation  Inflammatory response against extracellular pathogens Autoimmune response	Regulation of inflammatory response Regulation of autoimmune response  Suppression of anti-tumour immune response

\*The table summarises some of the main functions of the CD4<sup>+</sup> T cell subsets and does not imply that these are the only immune pathways and processes that they have been found to be implicated. A detailed description of each subset, signature cytokines and functions can be found in relevant publications<sup>18,19,75-78</sup>. IFN-γ: Interferon gamma; IL2-23: Interleukins 2-23; TGFβ: Transforming growth factor beta; T<sub>H</sub>1: T helper 1; T<sub>H</sub>2: T helper 2; T<sub>H</sub>9: T helper 9; T<sub>FH</sub>: T follicular helper; T<sub>H</sub>17: T helper 17; T<sub>Reg</sub>: Regulatory T cells.