WILEY

RESEARCH ARTICLE OPEN ACCESS

The Causal Role of Thyroid Hormones in Bipolar Disorders: A Two-Sample Mendelian Randomization Study

James L. Li[1,2] 📶

[1]Department of Public Health Sciences, University of Chicago, Chicago, Illinois, USA | [2]Interdisciplinary Scientist Training Program, University of Chicago, Chicago, Illinois, USA

Correspondence: James L. Li (jll1@uchicago.edu)

ABSTRACT

Introduction: Bipolar disorder is a complex psychiatric condition with distinctions between clinical subtypes including Type 1 and 2 disorders. Several studies have proposed that thyroid hormones may be involved in the aetiology of bipolar disorders.
Methods: This study employed a two-sample Mendelian randomization (MR) approach to investigate the causal relationships between six thyroid hormone metrics (TSH, FT4, FT3, TT3, FT3/FT4 and TT3/FT4) and bipolar disorder and Type 1 and 2 disorders, separately. Genome-wide association (GWAS) data from the ThyroidOmics Consortium (up to 271,040 individuals of European ancestry) were used for thyroid function metrics. Bipolar disorder GWAS data included 41,917 cases and 371,549 controls (25,060 Type 1 and 6,781 Type 2 cases). We applied inverse variance weighted (IVW) methods for primary MR analysis, with MR Egger, weighted median and weighted mode for sensitivity. Additional tests assessed horizontal pleiotropy and heterogeneity.
Results: Higher FT4 levels showed a protective causal effect against bipolar disorder (OR: 0.92, 95% CI: 0.86–0.97, $p = 4.58 \times 10^{-3}$) and a suggestive effect on Type 1 disorders (OR: 0.92, 95% CI: 0.86–0.99, $p = 3.21 \times 10^{-2}$). Elevated FT3 (OR: 1.18, 95% CI: 1.03–1.35, $p = 1.55 \times 10^{-2}$) and FT3/FT4 ratio (OR: 1.97, 95% CI: 1.02–3.82, $p = 4.46 \times 10^{-2}$) had suggestive harmful effects on Type 1 disorders. Sensitivity analyses showed consistent effects, with no significant horizontal pleiotropy or heterogeneity.
Conclusions: These findings highlight the protective role of FT4 and the potentially harmful effect of elevated FT3 in Type 1 bipolar disorder, highlighting the need for further research on thyroid hormone levels as a potential treatment strategy for Type 1 bipolar disorder.

1 | Introduction

Bipolar disorder is a complex psychiatric condition characterised by mood fluctuations that include episodes of mania or hypomania and depression [1]. The disorder is classified into two main subtypes: Type 1 bipolar disorder, which is marked by severe manic episodes, and Type 2, which involves less intense hypomanic episodes alongside depressive episodes [1]. Although the exact etiology of bipolar disorder remains unclear, some observational studies have identified biomarkers, including levels of thyroid hormones, that are associated with the disorder [2, 3]; however, because of possible confounding factors, observational studies cannot establish a definitive causal link between thyroid function and bipolar disorders.

Mendelian randomization (MR) has emerged as powerful tool for assessing the causal effects of exposures on outcomes by utilizing genetic variants as instruments [4]. While previous MR studies have suggested an association between free thyroxine (FT4) and overall bipolar disorder status [5, 6], the relationship

between other thyroid hormones and bipolar disorder subtypes remains unexplored, in part due to limited sample sizes used in earlier genome-wide association studies (GWAS) for thyroid functions. Furthermore, the direct causal relationship between thyroid hormones and specific subtypes of bipolar disorders has remained largely unexplored.

In this study, we conduct a two-sample MR analysis using recently published GWAS summary statistics [7, 8] involving individuals with European ancestry to assess the potential causal link between thyroid hormones (exposures) and bipolar disorders (outcomes). We specifically examine six thyroid hormone metrics: Thyroid-stimulating hormone (TSH), FT4, free triiodothyronine (FT3), total triiodothyronine (TT3), the ratio of FT3 to FT4 (FT3/FT4) and the ratio of TT3 to FT4 (TT3/FT4). Furthermore, we extend our analysis to investigate the causal relationships between these thyroid hormones and Type 1 and 2 bipolar disorders.

2 | Materials and Methods

2.1 | Obtaining GWAS Summary Statistics for Thyroid Hormones and Subtypes of Bipolar Disorder

GWAS summary statistics for TSH, FT4, FT3, TT3, FT3/FT4 and TT3/FT4 were obtained from Sterenborg et al. [7] and were based on 271,040, 119,120, 59,061, 15,829, 51,095 and 15,510 individuals with European ancestry, respectively, who were participants in the ThyroidOmics Consortium; these individuals were above 18 years of age, not using thyroid medications and did not have any history of thyroid surgeries. In generating these summary statistics, TSH, FT4, FT3 and TT3 were treated as continuous variables following an inverse normal transformation, and FT3/FT4 and TT3/FT4 underwent natural logarithm transformation [7]. Summary statistics for bipolar disorder, as well as Type 1 and 2 bipolar disorders, were obtained from separate studies detailed in Mullins et al. [8] involving a total of 41,917 bipolar disorder cases and 371,549 controls with European ancestry; of these cases, 25,060 individuals diagnosed with Type 1 bipolar disorder and 6781 individuals diagnosed with Type 2 bipolar disorder were included for generating subtype-specific summary statistics. Diagnoses of bipolar disorder in most of these cases were based on the international consensus criteria for lifetime bipolar disorder (DSM-IV, ICD-9, or ICD-10), which were determined through structured diagnostic interviews, clinician-administered checklists or medical record reviews [8]. All original studies contributing to these GWAS obtained ethical approval, and participants gave informed consent.

2.2 | Study Design of the Two-Sample Mendelian Randomization (MR)

For the causal estimates derived from a Mendelian randomization (MR) study to be considered valid, three key assumptions must be fulfilled: (1) the genetic variants must be strongly correlated with the exposure, (2) the genetic variants should not be linked to any confounders of the exposure–outcome relationship and (3) the variants must not influence the outcome independently of the exposure. In this study, we employed a two-sample MR approach utilizing GWAS summary statistics to assess the association between each thyroid hormone (exposures) on each bipolar subtype (outcomes), utilizing genetic variants as instrumental variables.

2.3 | Selection of Genetic Variants as Instruments

Separately for each thyroid hormone, we identified genetic variants associated with a $p < 5 \times 10^{-8}$ and performed linkage disequilibrium clumping using PLINK 1.9 [9] with a stringent threshold of $r^2 < 0.001$ and a 10,000 kb window; European individuals from the 1000 genomes project [10] were utilized as the reference panel for clumping. For thyroid hormones for which we identified very few genetic variants after clumping, we loosened the p-value threshold for clumping to 5×10^{-7} and 5×10^{-6} to obtain at least 10 initial genetic variants per thyroid hormone. Any genetic variants that were associated with the outcome ($p < 0.05$) were then excluded from MR analyses. To assess whether pleiotropic pathways confounded the MR results, we queried all instruments in the NHGRI-EBI Catalogue [11] to determine if any instruments were associated with potential confounders between thyroid function and bipolar disorders; variants which were associated with potential confounders (including psychiatric and neurological conditions, as well as properties related to neurobiology such as brain function) were excluded from our analyses. For the remaining variants, we calculated the F-statistics using the formula β^2/SE^2 to assess the strength of the selected genetic instruments, ensuring that all instruments used in our MR analyses had sufficient strength ($F > 10$). We determined the variance explained by a specific SNP using the formula $2\beta^2 \times f \times (1-f)$, where β represents the SNP effect on a given thyroid hormone metric and f represents the effect allele frequency. We aligned the GWAS summary statistics for both exposure and outcome by matching them based on the effect alleles and excluded palindromic variants with ambiguous allele frequencies from the MR analyses [12, 13].

2.4 | Performing Two-Sample MR Between Thyroid Hormone Metrics and Subtypes of Bipolar Disorder

The inverse-variance weighted (IVW) method with multiplicative random effects [14] was then applied to combine SNP-specific causal estimates for each bipolar disorder using the TwoSampleMR v0.6.7 [15] package in R; the IVW method is comparable to a two-stage least squares or allele score analysis when using individual-level data, and is, therefore, regarded as a conventional approach for MR. Given the interrelatedness of thyroid hormone metrics, we utilized a threshold of $p < 8.33 \times 10^{-3}$ from the IVW method to identify causal associations after accounting for multiple testing, corresponding to an alpha of 0.05 divided by six tests for the six thyroid hormone metrics. We reported suggestive causal associations using a threshold of $p < 0.05$.

2.5 | Sensitivity Analyses and Assessment of Pleiotropy

In addition to IVW, we utilized the TwoSampleMR package to implement MR-Egger [16], weighted median [17] and weighted

mode [18] MR methods to assess directional consistency of MR-IVW estimates. To confirm the robustness of significant findings, we conducted the Cochran's Q-statistic test for heterogeneity, as well as tests for nonzero MR-Egger intercepts [12] and the MR-Pleiotropy RESidual Sum and Outlier (MR-PRESSO) global test to assess pleiotropy [19]. For both causal ($p < 8.33 \times 10^{-3}$) and suggestive causal ($p < 0.05$) relationships we identified, we evaluated the influence of individual or pleiotropic variants by both computing MR estimates separately for each instrument, as well as conducting a leave-one-out analysis, where each SNP was systematically removed one at a time [16] as a sensitivity analysis. We also evaluated whether our study had sufficient power to detect the associations we observed between thyroid hormone metrics and bipolar disease by performing power calculations using the mRnd webtool [20]. Lastly, we assessed whether there was publication bias in the summary statistics utilized for determining MR instruments by creating funnel plots for the causal associations we identified.

3 | Results

3.1 | Genetic Variants Selected as Instruments for Each Thyroid Hormone Metric

We identified genetic variants for each thyroid hormone metric (FT3, FT4, TSH, TT3, FT3/FT4 and TT3/FT4) that met the assumptions to serve as instruments in MR analyses including a total of 52 variants for FT4, 152 for TSH, 12 for FT3/FT4 and less than 10 variants for either FT3, TT3 or TT3/FT4 at a p value threshold of 5×10^{-8} (Table S1). When relaxing the p value threshold to 5×10^{-7}, we identified 13 suitable instruments for FT3 and 11 for TT3/FT4, and further relaxing the threshold to 5×10^{-6} provided 11 variants that qualified as instruments for TT3 (Table S1). In total, we identified eight genetic variants among all instruments across the six thyroid hormone metrics that were associated with pleiotropic pathways in the NHGRI-EBI GWAS Catalogue [11] that may confound the MR results, including psychiatric and neurological conditions, as well as traits related to brain function/anatomy (Table S2); these eight variants are not included in the aforementioned numbers of instruments nor in any subsequent MR analyses. F-statistics for all valid instrumental variables were greater than 10, ranging from 20.84 to 2183.15, confirming the strength and reliability of these instruments for MR analyses (Table S1). Additional characteristics for the instruments utilized in MR analyses are presented in Table S1.

3.2 | Two-Sample Mendelian Randomization Demonstrates That Thyroid Hormones Causally Impact the Risk of Bipolar Disorder Subtypes

Utilizing genetic instruments for thyroid hormone metrics, we separately performed two-sample MR for overall bipolar disorder, as well as Type 1 and Type 2 bipolar disorders (Table S4). We identified a causal negative relationship between FT4 hormone levels and the odds of bipolar disorder, overall, using the IVW method (OR: 0.92, 95% CI: 0.86, 0.97, p value = 4.58×10^{-3}) even after adjusting for multiple testing (Table 1). The directionality of this

relationship between FT4 hormones and bipolar disorder identified by IVW was highly consistent across MR-Egger, weighted median and weighted mode MR approaches (Figure 1). We also identified that FT4 hormone levels had a suggestive causal negative relationship with Type 1 bipolar disorder (OR: 0.92, 95% CI: 0.86, 0.99, p value = 3.21×10^{-2}), which was also directionally consistent across other MR approaches (Table 1 and Figure 1). However, the MR association between FT4 and Type 2 bipolar disorder was not statistically significant (IVW p value = 0.29), and directionality of this relationship differed between MR approaches (Table S4). Furthermore, we identified two other thyroid hormone metrics, FT3 and FT3/FT4, that had suggestive causal associations with Type 1 bipolar disorder yielding IVW ORs of 1.18 (95% CI: 1.03, 1.35, p value = 1.55×10^{-2}) and 1.97 (95% CI: 1.02, 3.82, p value = 4.46×10^{-2}) respectively (Table 1); these associations were directionally consistent across MR methods (Table 1 and Figure 1). While the associations between FT3/FT4 and Type 2 bipolar disorder were not statistically significant ($p > 0.05$), the directionality of the associations seemed to be opposite to that of the association between FT3/FT4 and Type 1 bipolar disorder (Table S4). MR associations between FT3 and Type 2 bipolar disorder also appeared to be opposite in directionality compared to the association with Type 1 bipolar disorder, but these models showed evidence of directional pleiotropy (MR-Egger intercept p value = 0.048).

For the one causal and three suggestive causal relationships, we identified between thyroid hormone metrics and bipolar disorder subtypes (Table 1), we rigorously assessed potential biases by conducting several sensitivity analyses. The instruments utilized in these relationships did not show evidence of being in any confounding or pleiotropic pathways when examining genome-wide significant associations between these instruments and diseases/traits catalogued in NHGRI-EBI GWAS Catalogue (Table S4). We did not find evidence for pleiotropy across instrument effects (Cochran's Q test $p > 0.05$) (Table 1). Next, there was no evidence of directional pleiotropy among these causal associations (MR-Egger intercept $p > 0.05$) (Table 1). In addition, we computed the MR associations for each instrument separately and observed that they were generally consistent with the estimates inferred by each MR method collectively across all instruments (Figure 2). Moreover, MR-PRESSO did not detect any outlier genetic variants that needed to be removed (MR-PRESSO global test $p > 0.05$) (Table 1). To further corroborate the absence of outlier instruments, we performed a leave-one-out analysis where we computed MR estimates but left each instrument one by one; we observed that our findings were generally reliable and stable as none of the instruments heavily influenced the causal relationships we observed when excluded (Figure 3).

Additionally, we conducted power calculations for sample sizes of 413,466 (41,917 cases and 371,549 controls) for bipolar disorder overall and 396,609 (25,060 cases and 371,549 controls) for Type 1 bipolar disorder, with an alpha of 0.05. The sum of variance in FT4 explained by instruments in the MR with bipolar disorder overall was 3.84%, yielding a power of 0.87 to detect the MR association we observed. For Type 1 bipolar disorder, the sum of variance in FT4, FT3 and FT3/FT4 explained by instruments was 3.20%, 0.96% and 0.07%, respectively, yielding power values of 0.60, 0.76 and 0.97, respectively, for detecting the MR associations we observed.

TABLE 1 | | Causal and suggestive associations identified between thyroid hormone metrics and bipolar disorder subtypes in two-sample Mendelian randomization analyses.

Subtype	Thyroid hormone	MR method	Number of instruments	Odds ratio (95% CI)	p value	Cochran's Q p value	MR-egger Intercept p value	MR-PRESSO Global Test p value
Bipolar disorder (overall)	FT4	IVW	43	0.92 (0.86, 0.97)	4.58E-03*	0.81	0.69	0.83
		MR Egger	43	0.89 (0.78, 1.02)	0.11			
		Weighted median	43	0.89 (0.80, 0.98)	2.43E-02			
		Weighted mode	43	0.87 (0.79, 0.97)	1.30E-02			
Type 1 bipolar disorder	FT3	IVW	11	1.18 (1.03, 1.35)	1.55E-02**	0.79	0.75	0.85
		MR Egger	11	1.09 (0.68, 1.75)	0.72			
		Weighted median	11	1.15 (0.93, 1.43)	0.19			
		Weighted mode	11	1.14 (0.90, 1.45)	0.30			
	FT4	IVW	36	0.92 (0.86, 0.99)	3.21E-02**	0.95	0.80	0.96
		MR Egger	36	0.94 (0.79, 1.12)	0.49			
		Weighted median	36	0.90 (0.79, 1.04)	0.15			
		Weighted mode	36	0.90 (0.78, 1.03)	0.13			
	FT3/FT4	IVW	7	1.97 (1.02, 3.82)	4.46E-02**	0.32	0.41	0.46
		MR Egger	7	1.24 (0.37, 4.18)	0.74			
		Weighted median	7	1.61 (0.79, 3.28)	0.19			
		Weighted mode	7	1.62 (0.80, 3.27)	0.23			

Note: Odds ratios are based on FT4 and FT3 levels following an inverse normal transformation and FT3/FT4 undergoing a natural logarithm transformation.

*Causal association identified at multiple testing threshold (Bonferroni) of $p < 8.33 \times 10^{-3}$.

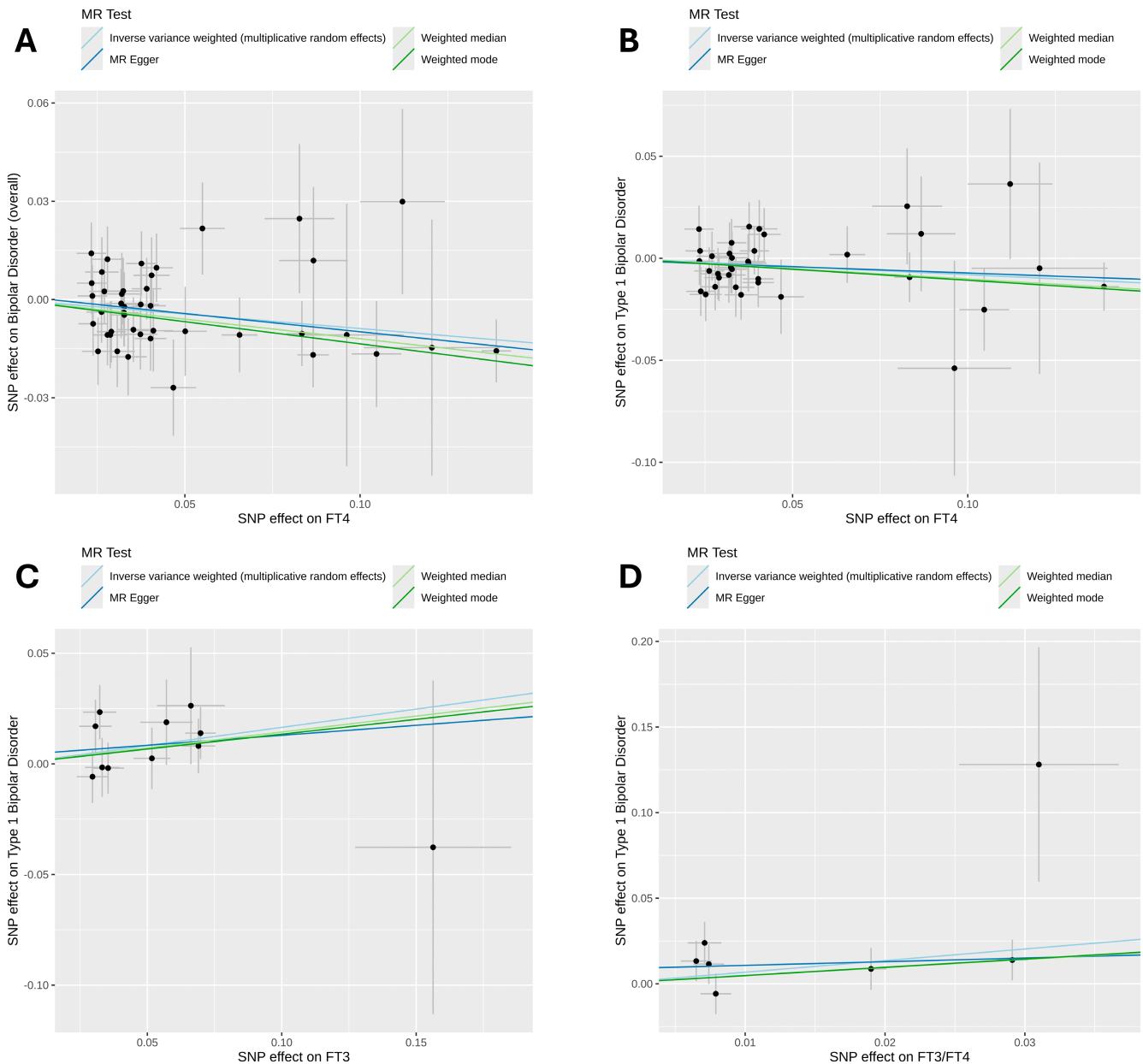**Suggestive causal association identified at a nominal significance threshold of $p < 0.05$.

FIGURE 1 | Scatterplots for the effects of instruments on each thyroid hormone metric and bipolar disorder subtype for causal and suggestive associations identified in Mendelian randomization (MR) analyses. (A) Relationship between FT4 and bipolar disorder (overall). (B) Relationship between FT4 and Type 1 bipolar disorder. (C) Relationship between FT3 and Type 1 bipolar disorder. (D) Relationship between FT3/FT4 and Type 1 bipolar disorder.

Funnel plots for each of the causal relationships we identified also suggested that there were no indications of publication bias in this study (Figure S1).

4 | Discussion

In this study, we conducted a two-sample MR of six thyroid hormone metrics and bipolar disorder overall, as well as Type 1 and 2 bipolar disorders. This study suggests that FT4 hormone levels may have a protective causal effect on bipolar disorder overall, with a suggestive causal effect on Type 1 bipolar disorder. We also identified suggestive harmful causal effects of elevated

FT3 levels and a higher FT3/FT4 ratio on Type 1 bipolar disorder. These findings remained consistent across various MR methods, each with different assumptions about horizontal and directional pleiotropy, suggesting that pleiotropy is unlikely to account for our results.

Our study aligns with previous Mendelian randomization research showing that FT4 has a protective impact on the risk of bipolar disorder [5, 6] and supports observational studies that found lower FT4 levels in bipolar disorder patients compared to controls [21]; however, our research specifically suggests that this relationship may be with Type 1 bipolar disorder. Although we cannot entirely dismiss the possibility that FT4 influences
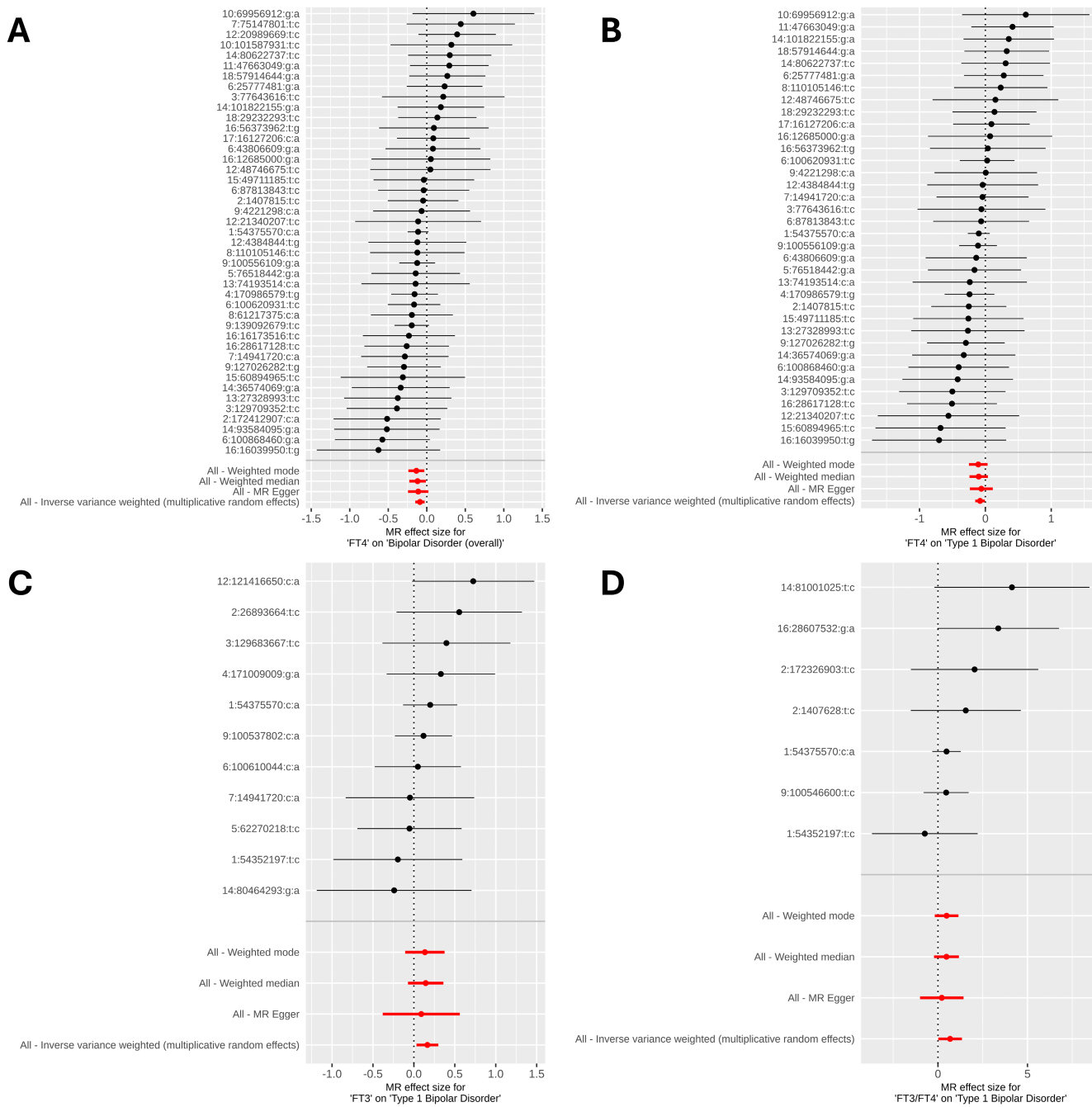
RIGHTSLINK()

FIGURE 2 | Forest plots of the Mendelian randomization (MR) estimates for each instrument individually compared to MR approaches using all instruments collectively. (A) Relationship between FT4 and bipolar disorder (overall). (B) Relationship between FT4 and Type 1 bipolar disorder. (C) Relationship between FT3 and Type 1 bipolar disorder. (D) Relationship between FT3/FT4 and Type 1 bipolar disorder.

Type 2 bipolar disorder, this uncertainty is largely due to the limited sample size of the Type 2 bipolar disorder GWAS summary statistics [8] used in our MR analysis. Additionally, our study reveals novel suggestive causal associations, showing a negative relationship between Type 1 bipolar disorder and both FT3 and the FT3/FT4 ratio. Notably, the opposing directionalities of the association between FT4 and Type 1 bipolar disorder and that between FT3, FT3/FT4 and Type 1 bipolar disorder are consistent with the conversion process of FT3 to FT4 in the body. This finding suggests that the conversion of FT3 to FT4 may, in part, have a protective effect against bipolar disorder; however, further research is needed to determine whether this

protective effect is solely due to increased FT4 production or if it is also influenced independently by decreased FT3 levels.

Furthermore, by validating previously reported MR associations between FT4 and bipolar disorder using much larger GWAS summary statistics for FT4 [7], with an approximately 65% increase in the number of participants compared to the latest prior study [5], our study provides more confidence that FT4 supplementation should be considered in the treatment of patients with Type 1 bipolar disorder. Interestingly, a recent clinical trial that administered levothyroxine and triiodothyronine as adjunctive treatments for patients who did not respond to lithium therapy
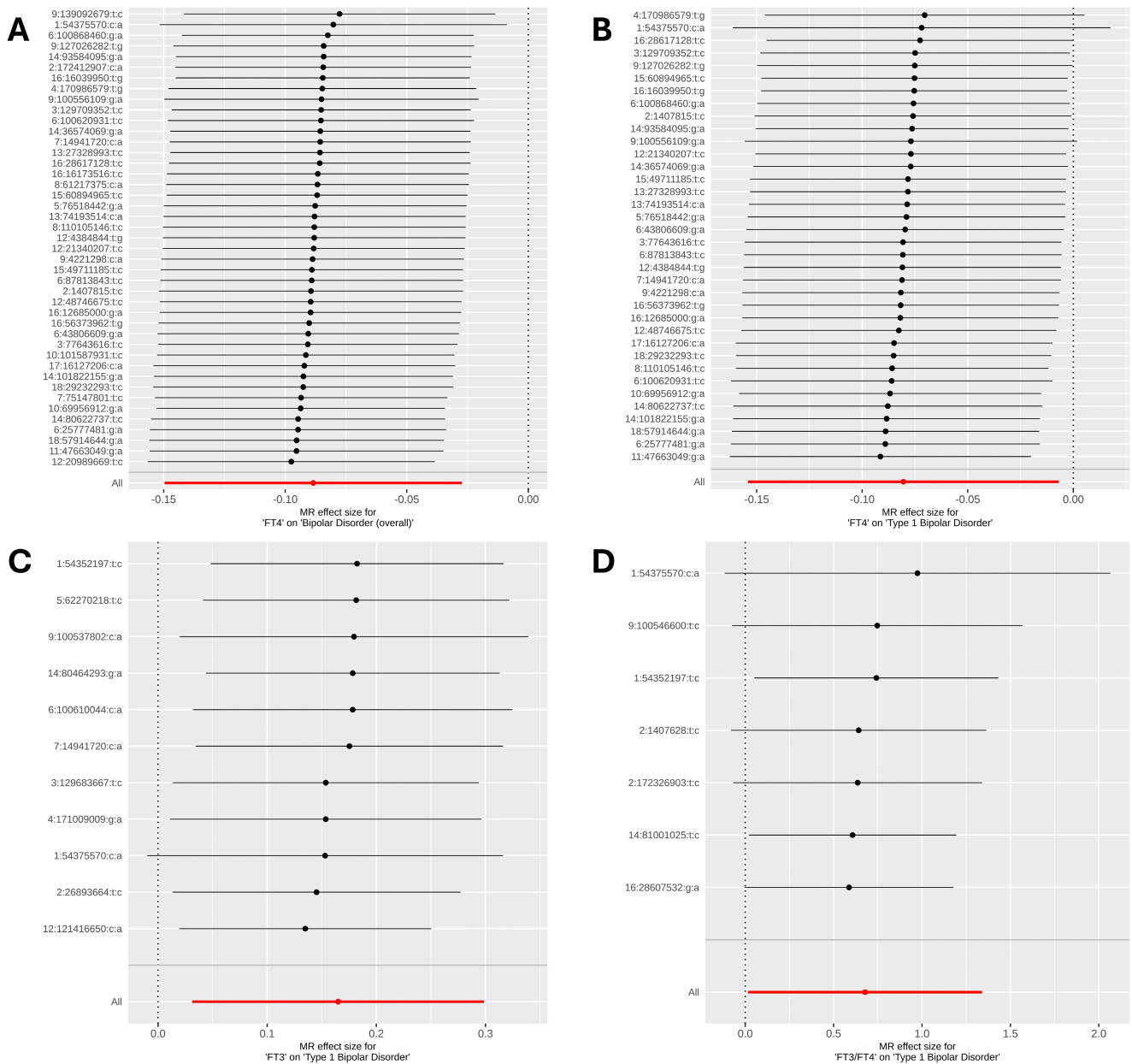
FIGURE 3 | Forest plots of the inverse-variance weighted Mendelian randomization (MR) method estimates utilizing the leave-one-out approach where instruments were excluded one by one. (A) Relationship between FT4 and bipolar disorder (overall). (B) Relationship between FT4 and Type 1 bipolar disorder. (C) Relationship between FT3 and Type 1 bipolar disorder. (D) Relationship between FT3/FT4 and Type 1 bipolar disorder.

demonstrated that levothyroxine had a positive effect on reducing mixed states [22]. Given that our results suggest that a reduction in FT3 levels may decrease the odds of Type 1 bipolar disorder, follow-up studies designed to explicitly evaluate the role of lowering FT3 while increasing FT4 levels are warranted.

Though this study managed to identify causal and suggestive causal relationships between thyroid functioning and bipolar disorders, there are several limitations. First, the sample size of individuals used to generate GWAS summary statistics for Type 2 bipolar disorder was quite underpowered [8] and may have potentially contributed to the lack of associations between thyroid hormone metrics and this subtype. Additionally, thyroid hormone measurements are quite variable even when taken from the same individual at different time points, including during

the daytime compared to in the evening [23]. Restricting participants included in the generation of thyroid function GWAS summary statistics to only those in a subset of studies with identical sample collection protocols (such as standardized time points for measurements) may help increase the precision of MR estimates. Furthermore, this study was constrained to individuals with European ancestry since GWAS summary statistics for both thyroid function metrics and bipolar disorder are quite limited in non-European populations; validation studies in other non-European populations are warranted.

Overall, these results contribute to a deeper understanding of the endocrine factors involved in bipolar disorders and underscore how both FT4 and FT3 may impact the risk of Type 1 bipolar disorder. These findings further suggest that thyroid

hormone metrics could potentially serve as biomarkers or therapeutic targets in the management of bipolar disorders.

Author Contributions

J.L.L. designed the study, performed data processing, conducted data analyses and wrote the manuscript.

Acknowledgements

Disclosure

The author affirms that the manuscript provides a truthful, accurate and transparent account of the reported study, with no significant aspects omitted.

Conflicts of Interest

The author declares no conflicts of interest.

Data Availability Statement

Data sharing is not applicable to this article as no new data were created or analyzed in this study. All GWAS summary statistics utilized in this study are previously published and publicly available and for academic use without restriction. GWAS summary statistics for thyroid function utilized in this study are downloadable from the ThyroidOmics Consortium at https://transfer.sysepi.medizin.uni-greifswald.de/thyroidomics/datasets/. GWAS summary statistics for bipolar disorders utilized in this study are downloadable from the Psychiatric Genomics Consortium at https://pgc.unc.edu/for-researchers/download-results/. Code for all analyses performed in this study is available at https://github.com/james-li-projects/MR_Thyroid_BPD.

References

1. I. Grande, M. Berk, B. Birmaher, and E. Vieta, "Bipolar disorder," *Lancet* 387, no. 10027 (2016): 1561–1572, https://doi.org/10.1016/S0140-6736(15)00241-X.

2. X. Song, Y. Feng, L. Yi, B. Zhong, and Y. Li, "Changes in Thyroid Function Levels in Female Patients With First-Episode Bipolar Disorder. Front," *Psychiatry* 14 (2023): 1185943, https://doi.org/10.3389/fpsyt.2023.1185943.

3. S. Chakrabarti, "Thyroid Functions and Bipolar Affective Disorder," *Journal of Thyroid Research* 2011 (2011): 306367, https://doi.org/10.4061/2011/306367.

4. D. A. Lawlor, R. M. Harbord, J. A. C. Sterne, N. Timpson, and S. G. Davey, "Mendelian Randomization: Using Genes as Instruments for Making Causal Inferences in Epidemiology," *Statistics in Medicine* 27, no. 8 (2008): 1133–1163, https://doi.org/10.1002/sim.3034.

5. G. Chen, H. Lv, X. Zhang, et al., "Assessment of the Relationships Between Genetic Determinants of Thyroid Functions and Bipolar Disorder: A Mendelian Randomization Study," *Journal of Affective Disorders* 298 (2022): 373–380, https://doi.org/10.1016/j.jad.2021.10.101.

6. Z. Gan, X. Wu, Z. Chen, et al., "Rapid Cycling Bipolar Disorder Is Associated With Antithyroid Antibodies, Instead of Thyroid Dysfunction," *BMC Psychiatry* 19, no. 1 (2019): 378, https://doi.org/10.1186/s12888-019-2354-6.

7. R. B. T. M. Sterenborg, I. Steinbrenner, Y. Li, et al., "Multi-Trait Analysis Characterizes the Genetics of Thyroid Function and Identifies Causal Associations With Clinical Implications," *Nature Communications* 15, no. 1 (2024): 888, https://doi.org/10.1038/s41467-024-44701-9.

8. N. Mullins, A. J. Forstner, K. S. O'Connell, et al., "Genome-Wide Association Study of More Than 40,000 Bipolar Disorder Cases Provides New Insights Into the Underlying Biology," *Nature Genetics* 53, no. 6 (2021): 817–829, https://doi.org/10.1038/s41588-021-00857-4.

9. S. Purcell, B. Neale, K. Todd-Brown, et al., "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses," *American Journal of Human Genetics* 81, no. 3 (2007): 559–575.

10. A. Auton, G. R. Abecasis, D. M. Altshuler, et al., "A Global Reference for Human Genetic Variation," *Nature* 526, no. 7571 (2015): 68–74, https://doi.org/10.1038/nature15393.

11. E. Sollis, A. Mosaku, A. Abid, et al., "The NHGRI-EBI GWAS Catalog: Knowledgebase and Deposition Resource," *Nucleic Acids Research* 51, no. D1 (2023): D977–D985, https://doi.org/10.1093/nar/gkac1010.

12. J. Bowden, G. Davey Smith, and S. Burgess, "Mendelian Randomization With Invalid Instruments: Effect Estimation and Bias Detection Through Egger Regression," *International Journal of Epidemiology* 44, no. 2 (2015): 512–525, https://doi.org/10.1093/ije/dyv080.

13. J. Bowden, M. F. Del Greco, C. Minelli, G. Davey Smith, N. Sheehan, and J. Thompson, "A Framework for the Investigation of Pleiotropy in Two-Sample Summary Data Mendelian Randomization," *Statistics in Medicine* 36, no. 11 (2017): 1783–1802, https://doi.org/10.1002/sim.7221.

14. S. Burgess, A. Butterworth, and S. G. Thompson, "Mendelian Randomization Analysis With Multiple Genetic Variants Using Summarized Data," *Genetic Epidemiology* 37, no. 7 (2013): 658–665, https://doi.org/10.1002/gepi.21758.

15. G. Hemani, J. Zheng, B. Elsworth, et al., "The MR-Base Platform Supports Systematic Causal Inference Across the Human Phenome," in *eLife*, ed. R. Loos, e34408 (Cambridge, UK: eLife Sciences Publications, 2018): 7, https://doi.org/10.7554/eLife.34408.

16. S. Burgess and S. G. Thompson, "Interpreting Findings From Mendelian Randomization Using the MR-Egger Method," *European Journal of Epidemiology* 32, no. 5 (2017): 377–389, https://doi.org/10.1007/s10654-017-0255-x.

17. J. Bowden, G. Davey Smith, P. C. Haycock, and S. Burgess, "Consistent Estimation in Mendelian Randomization With Some Invalid Instruments Using a Weighted Median Estimator," *Genetic Epidemiology* 40, no. 4 (2016): 304–314, https://doi.org/10.1002/gepi.21965.

18. F. P. Hartwig, G. Davey Smith, and J. Bowden, "Robust Inference in Summary Data Mendelian Randomization via the Zero Modal Pleiotropy Assumption," *International Journal of Epidemiology* 46, no. 6 (2017): 1985–1998, https://doi.org/10.1093/ije/dyx102.

19. M. Verbanck, C. Y. Chen, B. Neale, and R. Do, "Detection of Widespread Horizontal Pleiotropy in Causal Relationships Inferred From Mendelian Randomization Between Complex Traits and Diseases," *Nature Genetics* 50, no. 5 (2018): 693–698, https://doi.org/10.1038/s41588-018-0099-7.

20. M. J. A. Brion, K. Shakhbazov, and P. M. Visscher, "Calculating Statistical Power in Mendelian Randomization Studies," *International Journal of Epidemiology* 42, no. 5 (2013): 1497–1501, https://doi.org/10.1093/ije/dyt179.

21. T. S. J. Vedal, N. E. Steen, K. I. Birkeland, et al., "Free Thyroxine and Thyroid-Stimulating Hormone in Severe Mental Disorders: A Naturalistic Study With Focus on Antipsychotic Medication," *Journal of Psychiatric Research* 106 (2018): 74–81, https://doi.org/10.1016/j.jpsychires.2018.09.014.

22. P. D. Walshaw, L. Gyulai, M. Bauer, et al., "Adjunctive Thyroid Hormone Treatment in Rapid Cycling Bipolar Disorder: A Double-Blind Placebo-Controlled Trial of Levothyroxine (L-T4) and Triiodothyronine (T3)," *Bipolar Disorders* 20, no. 7 (2018): 594–603, https://doi.org/10.1111/bdi.12657.

23. S. Mahadevan, D. Sadacharan, S. Kannan, and A. Suryanarayanan, "Does Time of Sampling or Food Intake Alter Thyroid Function Test?," *Indian Journal of Endocrinology and Metabolism* 21, no. 3 (2017): 369–372, https://doi.org/10.4103/ijem.IJEM_15_17.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.