

1 *Molecular Psychiatry*

2

3 **MENTAL HEALTH TRAJECTORIES OF ADOLESCENTS TREATED WITH**
4 **PSYCHOTROPIC MEDICATIONS: INSIGHTS FROM THE EUROPEAN MILESTONE**
5 **STUDY**

6

7

8 Marta Magno^{1#}, Donato Martella^{1#}, Silvia Leone¹, Giovanni Allibrio², Angelo Bertani³, Elisa
9 Caselani¹, Patrizia Conti⁴, Samuele Cortese⁵, Gwen Dieleman⁶, Tomislav Franic⁷, Suzanne
10 Gerritsen⁶, Deborah Maffezzoni¹, Francesco Margari⁸, Ottaviano Martinelli⁹, Fiona McNicholas^{10 11}
11 ¹², Rocco Micciolo¹³, Renata Nacinovich^{14,15}, Diane Purper Ouakil^{16,17}, Adriana Pastore¹⁸, Francesco
12 Rinaldi¹⁹, Paramala Santosh²⁰, Paolo Scocco²¹, Ulrike Schulze^{22,23}, Swaran Singh⁺²⁴, Antonella
13 Squarcia²⁵, Paolo Stagi²⁶, Cathy Street²⁴, Elena Toffol¹, Helena Tuomainen²⁴, Larissa S van
14 Bodegom⁶, Stefano Vicari²⁷ & Giovanni de Girolamo¹ and on behalf the MILESTONE Consortium*

15

16 # Marta Magno and Donato Martella equally contributed to this work

17

18 **+ Corresponding author:**

19 Prof. Swaran P Singh,

20 Warwick Medical School, University of Warwick, United Kingdom

21 E-mail: S.P.Singh@warwick.ac.uk;

22

23 * A list of authors and their affiliations appears at the end of the paper

SUPPLEMENTARY MATERIAL

TABLE 1S.

SOCIODEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF THE EXCLUDED

SAMPLE (N = 314)

To evaluate the representativeness of the analytical sample (N = 690), we compared participants included in the analysis with those excluded due to missing medication data (N = 314). Overall, the two groups were generally comparable regarding demographic and clinical variables.

Demographic characteristics, such as age and ethnicity, were similar between the groups, with mean ages of 17.7 years for the included group and 17.4 years for the excluded group. Both samples predominantly consisted of individuals from a Caucasian background. However, the gender distribution differed slightly; the included group had a higher proportion of females (64.1%) compared to the excluded group (51.3%).

Clinical severity, as assessed by clinicians, was comparable across levels of illness severity. Mean scores for mental health symptoms, quality of life, and everyday functioning were also similar between the groups. Total HoNOSCA scores among the included groups ranged from 9.3 to 13.9, while the excluded group had a mean score of 13.8. The WHOQOL-BREF total scores for the included sample varied from 77.9 to 80.2, compared to a mean of 78.7 in the excluded group. Similarly, SLOF total scores for the included participants ranged from 183.5 to 192.5, whereas the excluded group had a mean score of 183.5.

The rates of lifetime suicide attempts and non-accidental self-injury did not significantly differ between groups, although the excluded group had slightly higher proportions of missing data for some clinical variables.

In conclusion, despite minor differences in gender distribution and rates of missing data, the included sample appears to be broadly representative of the full cohort.

	N (%) or mean (SD)
Gender	
<i>Female</i>	161 (51.3%)
<i>Male</i>	153 (48.7%)
Age	17.4 (0.31)
Ethnicity	
<i>Caucasian</i>	239 (76.1%)
<i>Other ethnic groups</i>	34 (10.8%)
<i>Missing</i>	41 (13.1%)
Living situation	
<i>With biological parents</i>	113 (36.0%)
<i>With one biological parent</i>	70 (22.3%)
<i>Adoptive/foster parent(s) or other living arrangements</i>	85 (27.1%)
<i>Missing</i>	46 (14.6%)
Current education	
<i>Secondary/vocational</i>	4 (1.3%)
<i>Higher (under/postgraduate)</i>	15 (4.8%)
<i>No current school attendance</i>	170 (54.1%)
<i>Other</i>	7 (2.2%)
<i>Missing</i>	118 (37.6%)
Clinician rated severity of psychopathology (CGI-S)	
<i>Not at all ill</i>	27 (8.6%)

<i>Borderline/mildly/moderately ill</i>	164 (52.2%)
<i>Markedly ill or more severe</i>	65 (20.7%)
<i>Missing</i>	58 (18.5%)
Mental health (HoNOSCA; range 0–52)	13.8 (10.3)
Lifetime suicide attempt	
<i>Yes</i>	83 (24.6%)
<i>No</i>	178 (56.7%)
<i>Missing</i>	53 (16.9%)
Non-accidental self-injury (HoNOSCA domain)	
<i>No problem of this kind</i>	167 (53.2%)
<i>Occasional thoughts about death, or of self-harm not leading to injury. No self-harm or suicidal thoughts.</i>	23 (7.3%)
<i>Non-hazardous self-harm whether or not associated with suicidal thoughts</i>	25 (8.0%)
<i>Moderately severe suicidal intent or moderate non-hazardous self-harm</i>	19 (6.1%)
<i>Serious suicidal attempt or serious deliberate self-injury</i>	12 (3.8%)
<i>Missing</i>	68 (21.6%)
Quality of life (WHOQOL-BREF; range 4–20)	
<i>Psychological</i>	18.2 (3.8)
<i>Physical</i>	20.7 (3.7)
<i>Social</i>	10.3 (2.7)
<i>Environmental</i>	29.5 (5.4)
<i>Total</i>	78.7 (12.8)
Everyday functional skills (SLOF)	
<i>Physical functioning</i>	24.1 (1.4)
<i>Personal care skills</i>	32.6 (3.7)
<i>Interpersonal relationships</i>	25.8 (7.0)
<i>Social acceptability</i>	30.9 (3.5)
<i>Activities</i>	47.9 (8.7)
<i>Work skills</i>	22.1 (6.5%)
<i>Total</i>	183.5 (22.9)
Life events (range 0–13)	1.8 (1.6)
Bullying	
<i>Victim</i>	72 (22.9%)
<i>Bully/victim</i>	92 (29.3%)
<i>Bully</i>	17 (5.4%)
<i>Non-involved</i>	38 (12.1%)
<i>Missing</i>	95 (30.3%)

Table 2S.
LIST OF STANDARDISED TOOLS USED

TOOLS	DESCRIPTION	RANGE	ADMINISTERED TO
Achenbach System of Empirically Based Assessment (ASEBA)	The Achenbach System of Empirically-Based Assessment [ASEBA] have been used extensively in different contexts to evaluate emotional and behavioural problems in the last 6 months.	Higher scores indicate more emotional/behavioural problems	Patients and parents
Clinical Global Impression – Severity scale (CGI-S)	Instrument to measure the severity of the disorder	1-7	Clinicians
Health of the Nation Outcome Scale for Children and Adolescents (HoNOSCA)	The Health of the Nation Outcome Scale for Children and Adolescents [HoNOSCA] measures general health and social functioning in child and adolescent attending mental health services	0-52	Patients, parents and clinicians
World Health Organization Quality of Life Brief Inventory (WHOQOL-BREF)	Self-administered questionnaire to assess subjective health and well-being over the previous two weeks. It covers four domains: physical health, psychological health, social relationships, and environment.	0-100	Patients
Specific Levels of Functioning (SLOF)	To assess psychosocial functioning in six subscales: physical functioning, self-care skills, interpersonal relationships, social acceptance, community life activities, and work skills.	43-125	Parents
Life events	Instrument developed specifically for this project to assess life events	0-13	Patients
Retrospective Bullying and Friendship Interview Schedule (adapted version)	Instrument used to assess the patient experiences with bullying in different settings	NA	Patients

TABLE 3S.

CORRELATION BETWEEN SELF-REPORTED AND OBSERVER-REPORTED ASSESSMENTS

	Correlation Coefficient	CI (95%)	p-value
HoNOSCA Total Rating - Patients vs. Clinicians			
Time Point 1	0.508	0.451 - 0.561	< 0.001
Time Point 2	0.608	0.559 - 0.653	< 0.001
Time Point 3	0.599	0.549 - 0.645	< 0.001
Time Point 4	0.626	0.578 - 0.669	< 0.001
Overall	0.589	0.564 - 0.613	< 0.001
Internalising Factors: Self-Report (ASR+YSR) vs. Parent-report (CBCL+ABCL)			
Time Point 1	0.576	0.523 - 0.623	< 0.001
Time Point 2	0.595	0.544 - 0.641	< 0.001
Time Point 3	0.616	0.567 - 0.660	< 0.001
Time Point 4	0.635	0.588 - 0.677	< 0.001
Overall	0.605	0.581 - 0.628	< 0.001
Externalising Factors: Self-Report (ASR+YSR) vs. Parent-report (CBCL+ABCL)			
Time Point 1	0.563	0.509 - 0.612	< 0.001
Time Point 2	0.535	0.479 - 0.587	< 0.001
Time Point 3	0.586	0.534 - 0.633	< 0.001
Time Point 4	0.511	0.453 - 0.564	< 0.001
Overall	0.545	0.519 - 0.571	< 0.001

TABLE 4S.
DISTRIBUTION OF DIAGNOSES ACROSS MEDICATION GROUPS AT BASELINE

	Group 1* (N = 186)	Group 2* (N = 304)	Group 3* (N = 200)	Total sample (N = 690)
Anxiety	28 (15.1%)	38 (12.5%)	27 (13.5%)	93 (13.4%)
Depression	42 (22.6%)	62 (20.4%)	29 (14.5%)	133 (19.3%)
Neurodevelopmental	69 (37.1%)	85 (28.0%)	75 (37.5%)	229 (33.2%)
OCD/ED	19 (10.2%)	34 (11.2%)	23 (11.5%)	76 (11.0%)
Other	24 (12.8%)	63 (20.7%)	37 (18.5%)	124 (18.0%)
Missing	4 (2.2%)	22 (7.2%)	9 (4.5%)	35 (5.1%)
<p>* Group 1 = Continuous Medication, Group 2 = Intermittent Medication, Group 3 = Never Medicated The diagnostic category 'Other' comprises conditions including personality disorders, conduct disorders, substance use disorders and schizophrenia</p>				

TABLE 5S.

NUMBER OF MEDICATIONS TAKEN AT EACH TIME POINT IN MEDICATED PATIENTS

Medication (n.)	T1	T2	T3	T4
<i>Continuous Medication (N = 186)</i>				
1	107 (57.3%)	106 (57.0%)	95 (51.1%)	111 (59.7%)
2	60 (32.3%)	57 (30.6%)	67 (36.0%)	57 (30.6%)
3	14 (7.5%)	18 (9.7%)	14 (7.5%)	16 (8.6%)
4	4 (2.2%)	3 (1.6%)	8 (4.3%)	0 (0.0%)
5	1 (0.5%)	2 (1.1%)	2 (1.1%)	2 (1.1%)
<i>Intermittent Medication (N = 304)</i>				
0	99 (32.6%)	174 (57.2%)	185 (60.9%)	201 (66.1%)
1	159 (52.3%)	99 (32.6%)	100 (32.9%)	79 (26.0%)
2	37 (12.1%)	19 (6.2%)	15 (4.8%)	17 (5.6%)
3	7 (2.3%)	8 (2.7%)	2 (0.7%)	6 (2.0%)
4	2 (0.7%)	3 (1.0%)	2 (0.7%)	0 (0.0%)
5	0 (0.0%)	1 (0.3%)	0 (0.0%)	1 (0.3%)

TABLE 6S.

PATTERN OF PRESCRIPTIONS IN THE MEDICATED SAMPLE AT EACH TIME POINT

	T1	T2	T3	T4
<i>Continuous Medication (N = 186)</i>				
First-Generation Antipsychotic	8 (4.3%)	8 (4.3%)	9 (4.8%)	12 (6.5%)
Second-Generation Antipsychotic**	59 (31.7%)	61 (32.8%)	65 (34.9%)	60 (32.3%)
Mood Stabilizers	19 (10.2%)	20 (10.8%)	20 (10.8%)	20 (10.8%)
Antidepressants	119 (63.9%)	112 (60.2%)	123 (66.1%)	108 (58.1%)
Benzodiazepines	17 (9.1%)	18 (9.7%)	15 (8.1%)	8 (4.3%)
Stimulants	60 (32.3%)	63 (33.9%)	62 (33.3%)	60 (32.3%)
<i>Intermittent Medication (N = 304)</i>				
First-Generation Antipsychotic	8 (2.6%)	3 (1.0%)	3 (1.0%)	2 (0.7%)
Second-Generation Antipsychotic**	47 (15.5%)	25 (8.2%)	20 (6.6%)	11 (3.6%)
Mood Stabilizers	7 (2.3%)	4 (1.3%)	6 (2.0%)	8 (2.6%)
Antidepressants	109 (35.9%)	72 (23.7%)	46 (15.1%)	36 (11.8%)
Benzodiazepines	30 (9.9%)	10 (3.3%)	12 (3.9%)	16 (5.3%)
Stimulants	44 (14.5%)	35 (11.5%)	19 (6.2%)	10 (3.3%)

FIGURE S1

FLOWCHART OF THE STUDY

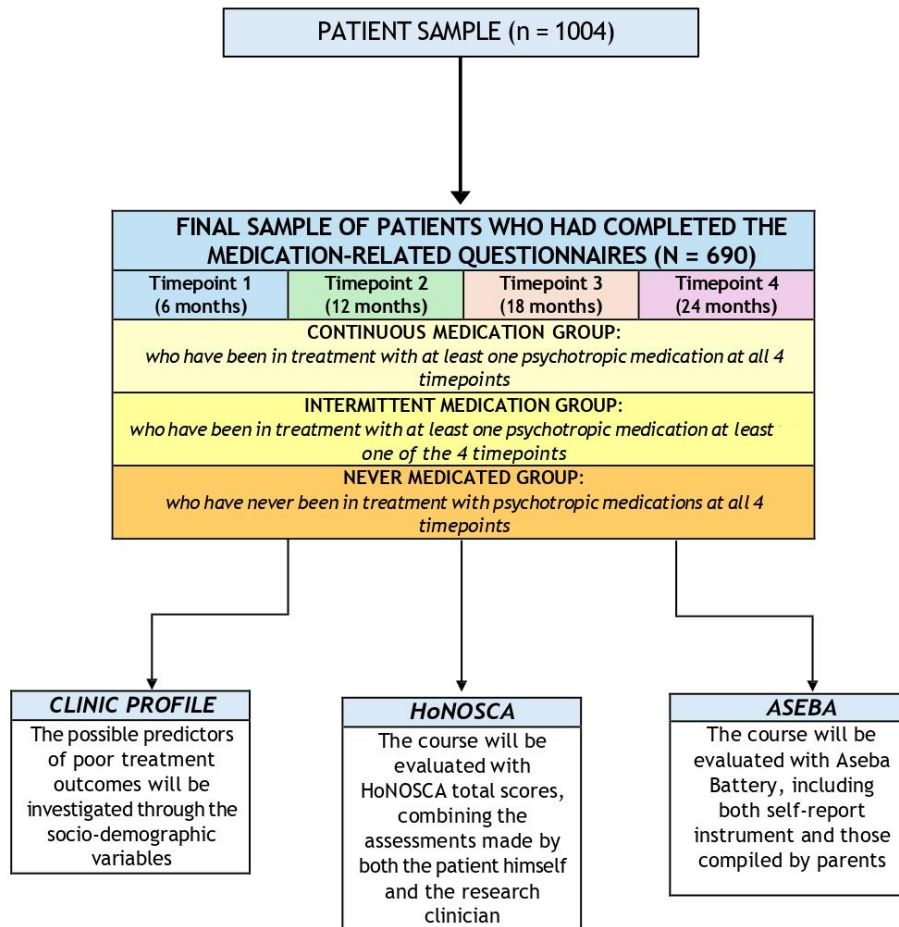


FIGURE S2.

HONOSCA TOTAL RATING - PATIENTS VS. CLINICIANS

Figure S2 provides a general view of patients' courses over 2 years, and shows the HoNOSCA total ratings, comparing patient and clinician ratings across four time points (T1 to T4). Both groups show a general decline in ratings, indicating an overall reduction in symptom severity over time. However, the pattern of decrease differs between the two. Patients reported a more marked improvement in symptoms, particularly noticeable between T1 and T4 ($p < 0.001$), with a significant overall reduction ($p < 0.001$). In contrast, clinicians reported higher severity at T1 compared to patient self-reports and noted a gradual, linear, and significant ($p < 0.001$) decrease in symptoms, which was less pronounced than that reported by patients. For these trends, there is statistically significant difference between the two ratings at any of the four time points ($p = 0.042$).

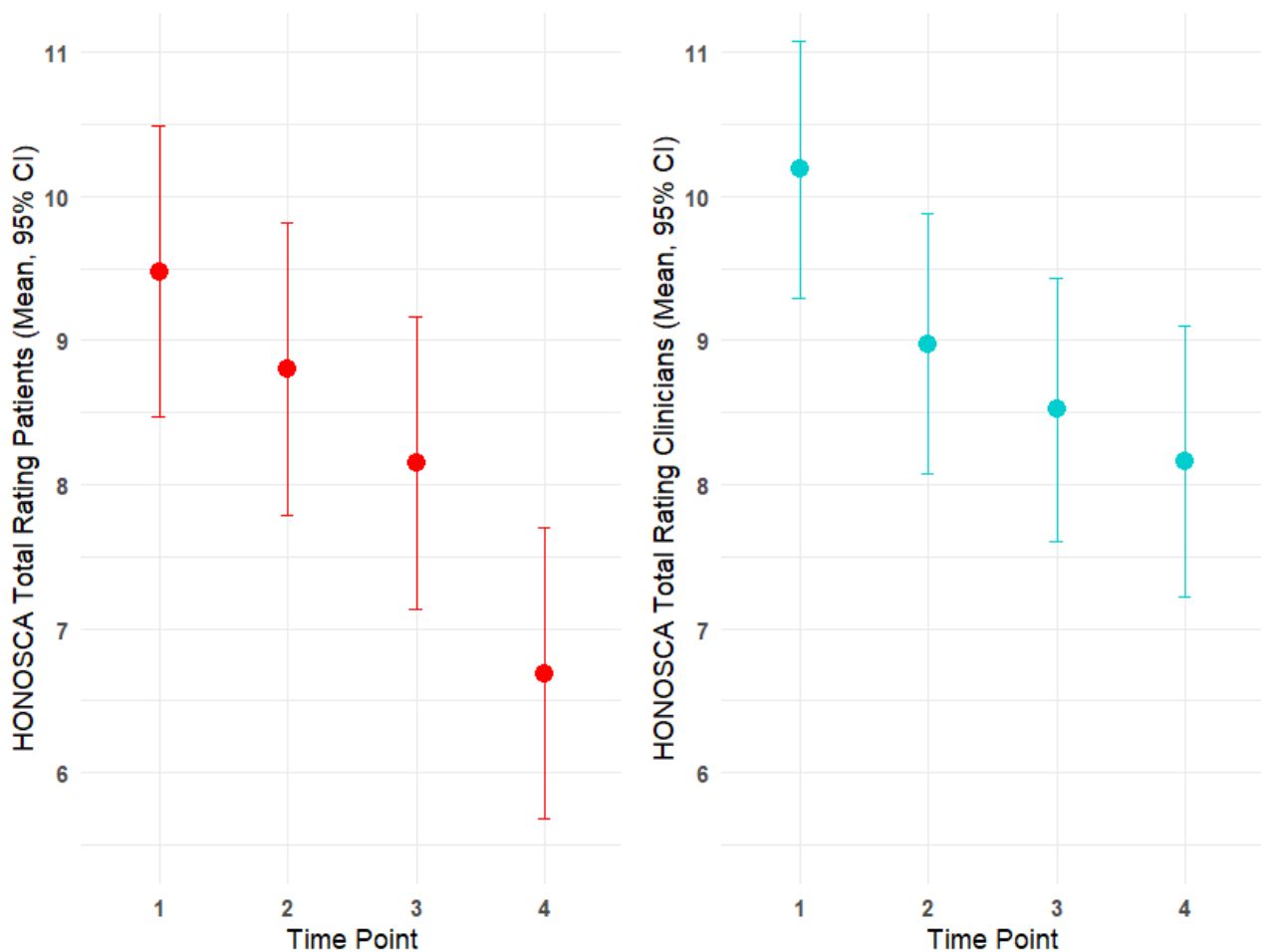


FIGURE S3.

INTERNALISING FACTORS ANALYSIS: SELF-REPORT (ASR+YSR) VS. PARENT-REPORT (CBCL+ABCL)

Figure S3 presents the analysis of ASEBA internalising scales, comparing self-reported (YSR and ASR ratings) and parent-reported (CBCL and ABCL scores) assessments over four time points. There was a significant decrease in ASR and YSR ratings, suggesting an improvement in internalising domains as reported by the subjects themselves ($p < 0.001$). The downward trend observed in the CBCL and ABCL ratings, while indicating a reduction, was statistically significant ($p < 0.001$). For these trends, there was a statistically significant difference between the self-report and parent-report scales at all four time points ($p = 0.017$), highlighting a disparity in perceived internalising symptoms between the adolescents and parent/carers.

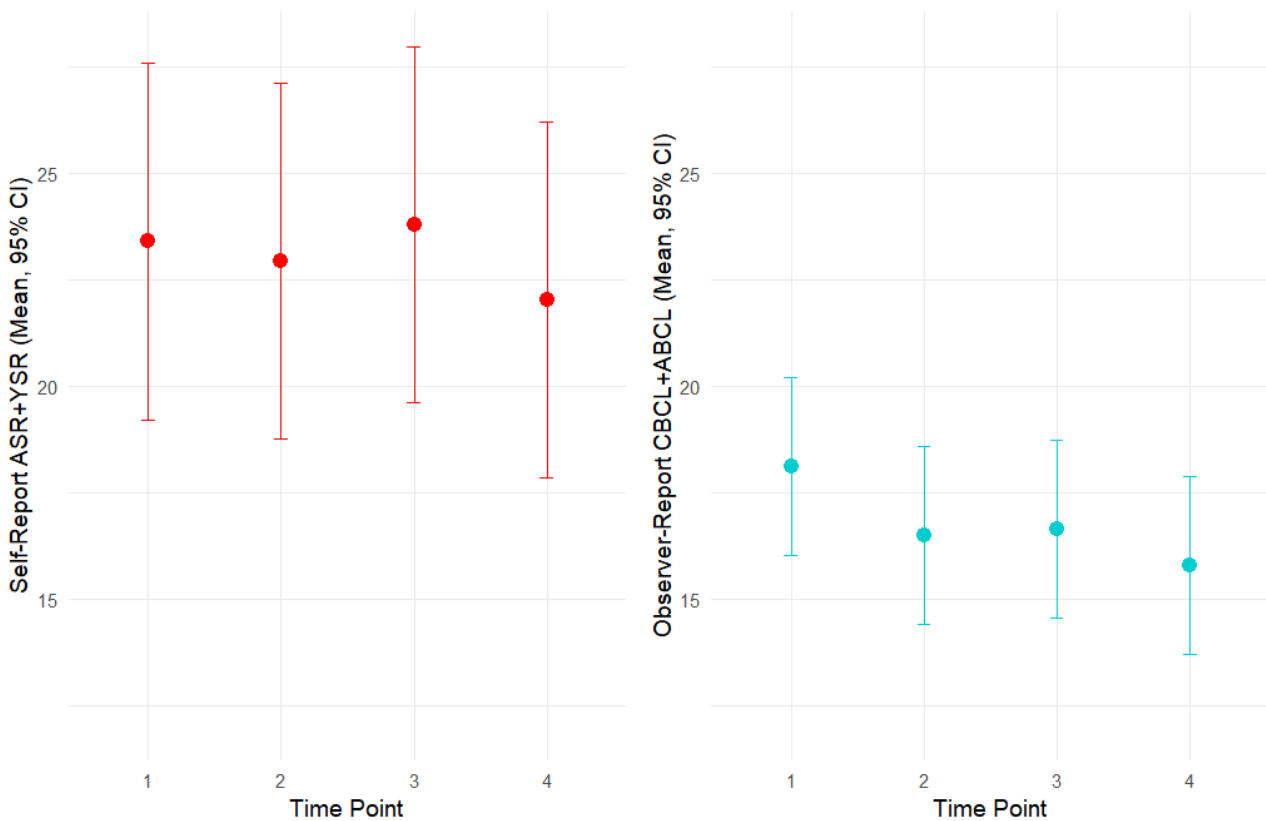


FIGURE S4.

EXTERNALISING FACTORS ANALYSIS: SELF-REPORT (ASR+YSR) VS. PARENT-REPORT (CBCL+ABCL)

Figure S4 illustrates the analysis of ASEBA externalising scales, comparing self-reported (YSR and ASR ratings) and parent-reported (CBCL and ABCL scores) assessments. The results showed a noticeable decrease in the ASR and YSR ratings over time, with average scores significantly diminishing ($p < 0.001$), suggesting a potential improvement in externalising behaviours as self-reported by the subjects. In contrast, the CBCL and ABCL ratings, which are based on parents'/carers' reports, did not exhibit a significant change over time ($p = 0.157$). Additionally, when comparing the changes over time between the self-reports and parents'/carers' reports, there was statistically significant difference between the two scales at any of the time points ($p < 0.001$).

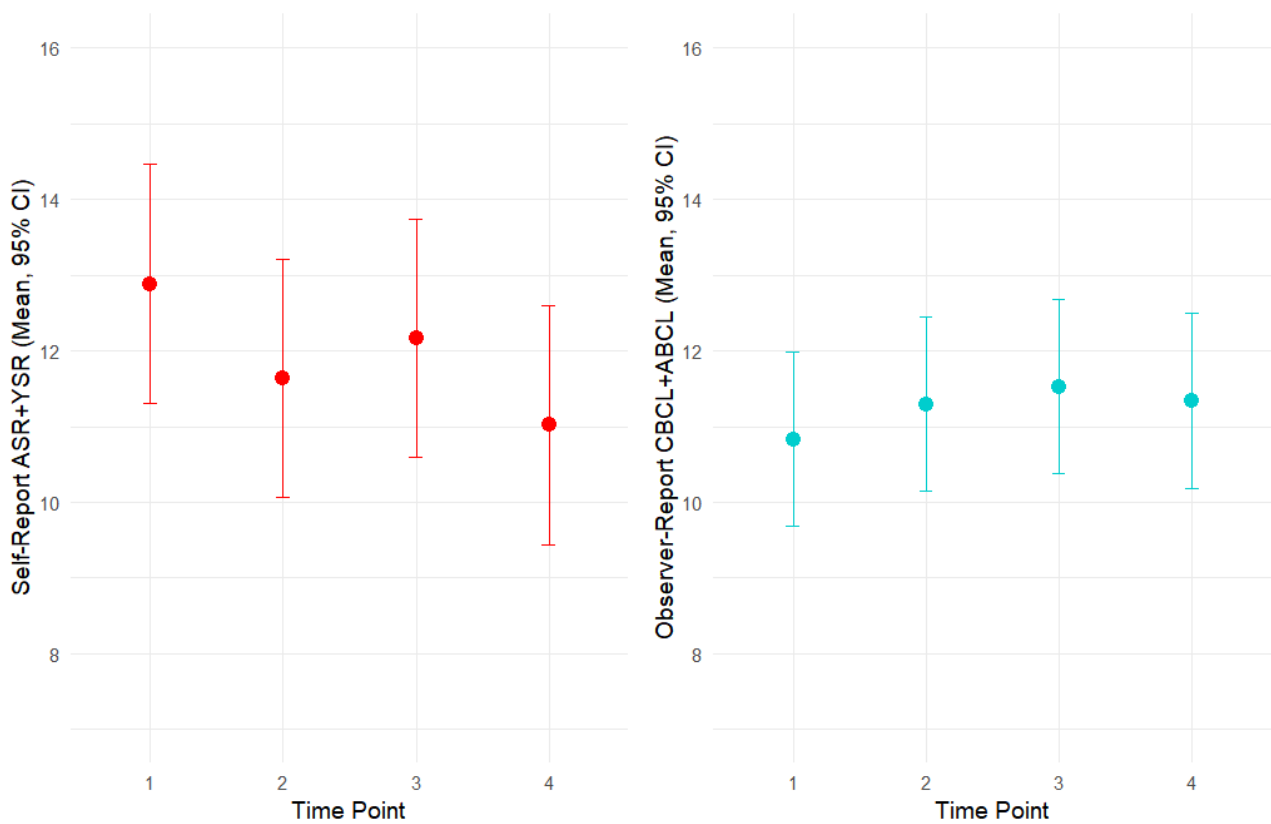


FIGURE S5
NUMBER OF MEDICATIONS TAKEN AT EACH TIME POINT IN PATIENTS WITH
CONTINUOUS MEDICATION ADHERENCE (N = 186)

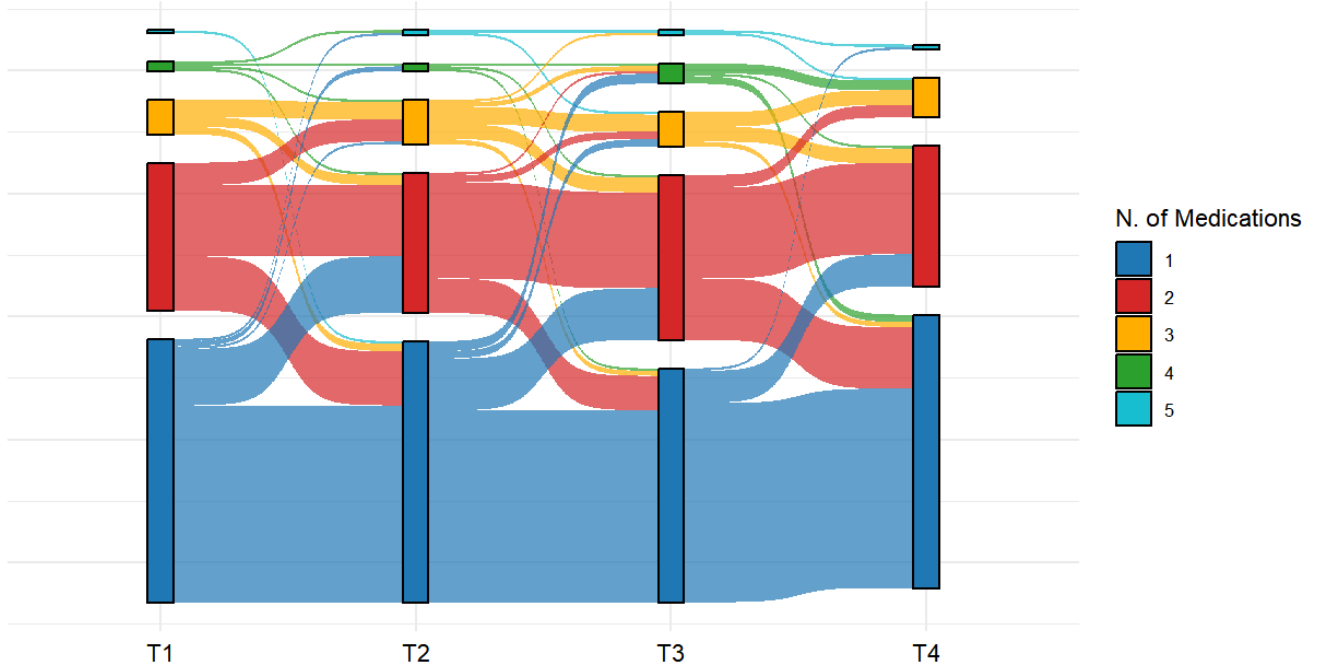


FIGURE S6
NUMBER OF MEDICATIONS TAKEN AT EACH TIME POINT IN PATIENTS WITH
INTERMITTENT MEDICATION ADHERENCE (N = 304)

