Supplementary material

Mapping trajectories of child and adolescent psychopathology: universal early identification of mental health needs

Ida Scheel Rasmussen, Philip Wilson, Gritt Overbeck, Claus Thorn Ekstrøm, Else Marie Olsen,

Katrine Strandberg-Larsen

Table S1: ICD codes for major congenital anomalies, conditions affecting the central nervous system and learning disabilities.

Action diagnoses	International Classification of Diseases (ICD) - 10
Major congenital anomalies diagnosed <1 year of age*	Q00, Q01, Q02, Q03, Q04, Q05, Q06, Q07, Q10-Q18, Q20-Q26, Q300, Q32- Q34, Q35-Q37, Q38-Q45, Q790, Q792, Q793, Q795, Q60-Q64, Q794, Q65-Q74, Q65-Q74, Q7402, Q77, Q7800, Q782-Q788, Q750, Q893, Q894, Q80-Q82, 8726, Q0435, Q411, Q412, Q418, Q710, Q712, Q713, Q720, Q722, Q723, Q730, Q793, Q795, Q7980, Q7982, Q8706, Q206, Q240, Q3381, Q890, Q893, Q86, P350, P351, P371, Q860, Q8680, P350, P351, P371, Q4471, Q6190, Q7484, Q751, Q754, Q7581, Q87, Q936, D821, Q90-Q92, Q93, Q96- Q99
Conditions affecting the central nervous system diagnosed <7 year of age	G10-G14, G35-G37, G40, G80-G83, G91
Learning disabilities	F70, F71, F72, F73, F78, F79

* The EUROCAT classification guided the definition of the major congenital anomalies (EUROCAT Guide 1.4: Instruction for the registration of congenital anomalies. EUROCAT Central Registry, University of Ulster, 2013)

Table S2: Dimensions and definitions of child and adolescent psychopathology

	Coverage	Definition	Data origin	Time points	Type
Behavioural problems					
1	Internalizing problem score	0-20	DNBC	7 (parent), 11 (teacher, parent or self-reported), 18 (self-reported) years of age	Psychometric scale
2	Externalizing problem score	0-20	DNBC	7 (parent), 11 (teacher, parent or self-reported), 18 (self-reported) years of age	Psychometric scale
Psychiatric diagnosis					
3	Neurodevelopmental diagnoses	1 count per unique psychiatric diagnoses within each year	LPR	Yearly from age 1 to age 18	Zeroinflated Poisson
4	Affective diagnoses	1 count per unique psychiatric diagnoses within each year	LPR	Yearly from age 1 to age 18	Zeroinflated Poisson
5	Other psychiatric diagnoses	1 count per unique psychiatric diagnoses calculated for each year	LPR	Yearly from age 1 to age 18	Zeroinflated Poisson
Redeemed prescriptions					
6	Psychotropic medicine	1 count pr. year with min. 1 redeemed prescription of psychoactive medication	LMDB	Yearly from age 1 to age 18	Binary

DNBC: Danish National Birth Cohort, LPR: The National Patient Register, LMDB: The Danish National Prescription Registry

Table S3: Inclusion and division of diagnoses in the dimension of psychiatric diagnosis

Dimension	ICD-10 code
Neurodevelopmental diagnoses[1,2,3]	F80-83 (other developmental disorder), F84 (autism spectrum disorder), F88-89 (other/unspecified disorders of psychological development), F90 (ADHD), F95 (tic disorders), F98.8 (ADHD – inattentive type), F20-F29 (schizophrenia spectrum disorder)

Supplementary material

Dimension	ICD-10 code
Affective disorders	F30-39 (mood disorders), F40-48, F93 (anxiety disorders)
Other psychiatric diagnoses	F50 (eating disorder), F51 (sleep disorders), F91 (ODD/CD), F94 (attachment disorder), F98.0-98.7, F98.9 (other behavioural and emotional disorders), F99 (psychiatric disorder unspecified), F60-F69 (personality disorders)

[1]Thapar A, Cooper M, Rutter M. Neurodevelopmental disorders. Lancet psychiatry. 2017 Apr;4(4):339-346. doi: 10.1016/S2215-0366(16)30376-5.

[2] Insel, T. R. (2010). Rethinking schizophrenia. Nature, 468(7321), 187-193.

[3] American Psychiatric Association, DSM-5 Task Force. (2013). Diagnostic and statistical manual of mental disorders: DSM-5™ (5th ed.). American Psychiatric Publishing, Inc.. https://doi.org/10.1176/appi.books.9780890425596.

Table S4: Prescriptions included in the dimension of psychoactive medication

ATC code	Name	Prescribed for
N06B	Psychostimulants	Agents used for ADHD, and Nootropics
N06A	Antidepressants	Agents used for depression
N05B	Anxiolytics	Agents used for severe anxiety
N05A	Antipsychotics	Agents used to manage psychosis
N05CH	Melatonin	Agent used as adjuvant treatment in affective disorders, ADHD and in schizophrenia[1]. Melatonin is also used for sleep problems[1].

[1] Bliddal, M., Kildegaard, H., Rasmussen, L. et al. Melatonin use among children, adolescents, and young adults: a Danish nationwide drug utilization study. Eur Child Adolesc Psychiatry 32, 2021–2029 (2023). https://doi.org/10.1007/s00787-022-02035-1.

Table S5: Hypothetical example of a study participant

Dimension		Psychopathology	Age							
			1-2	[…]	7-8	[…]	11-12	[…]	16-17	17-18
1	Internalizing problem score	SDQ-score	..		12		15		..	14
	Sum for dimension 1		**..**		**12**		**15**		**..**	**14**
2	Externalizing problem score	SDQ-score	..		7		8		..	6
	Sum for dimension 2		**..**		**7**		**8**		**..**	**6**
3	Neurodevelopmental diagnoses	F80 Other developmental disorder	0		0		0		0	0
		F81 Other developmental disorder	0		0		0		0	0
		F82 Other developmental disorder	0		0		0		0	0
		F83 Other developmental disorder	0		0		0		0	0
		F84 Autism spectrum disorder	0		0		1		0	0
		F88 Other/unspecified disorders of psychological development	0		0		0		0	0
		F89 Other/unspecified disorders of psychological development	0		0		0		0	0
		F90 ADHD	0		0		0		0	0
		F95 Tic disorder	0		0		0		0	0
		F98.8 ADHD – inattentive type	0		0		0		0	0
		F20 Schizophrenia spectrum disorder	0		0		0		0	0
		F21 Schizophrenia spectrum disorder	0		0		0		0	0
		F22 Schizophrenia spectrum disorder	0		0		0		0	0

Supplementary material

Dimension		Psychopathology	Age							
		F23 Schizophrenia spectrum disorder	0		0		0		0	0
		F24 Schizophrenia spectrum disorder	0		0		0		0	0
		F25 Schizophrenia spectrum disorder	0		0		0		0	0
		F26 Schizophrenia spectrum disorder	0		0		0		0	0
		F27 Schizophrenia spectrum disorder	0		0		0		0	0
		F28 Schizophrenia spectrum disorder	0		0		0		0	0
		F29 Schizophrenia spectrum disorder	0		0		0		0	0
	Sum for dimension 3		**0**		**0**		**1**		**0**	**0**
4	Affective diagnoses	F30 mood disorders	0		0		0		0	0
		F31 mood disorders	0		0		0		0	0
		F32 mood disorders	0		0		0		1	0
		F33 mood disorders	0		0		0		0	0
		F34 mood disorders	0		0		0		0	0
		F35 mood disorders	0		0		0		0	0
		F36 mood disorders	0		0		0		0	0
		F37 mood disorders	0		0		0		0	0
		F38 mood disorders	0		0		0		0	0
		F39 mood disorders	0		0		0		0	0
		F40 Anxiety	0		0		0		1	1
		F41 Anxiety	0		1		0		0	0
		F42 Anxiety	0		0		0		0	0
		F43 Anxiety	0		0		0		0	0
		F44 Anxiety	0		0		0		0	0
		F45 Anxiety	0		0		0		0	0
		F46 Anxiety	0		0		0		0	0
		F47 Anxiety	0		0		0		0	0
		F48 Anxiety	0		0		0		0	0
		F93 Anxiety	0		0		0		0	0
	Sum for dimension 4		**0**		**1**		**0**		**2**	**1**
5	Other psychiatric diagnoses	F50 Eating disorder	0		0		0		0	0
		F51 sleep disorders	0		0		0		0	0
		F91 ODD/CD	0		0		0		0	0
		F94 attachment disorder	0		0		0		0	0
		F98.0-98.7 Other behavioural and emotional disorders	0		0		0		0	0
		F99 (psychiatric disorder unspecified),	0		0		0		0	0
		F60 Personality disorders	0		0		0		0	0
		F61 Personality disorders	0		0		0		0	0
		F62 Personality disorders	0		0		0		0	0
		F63 Personality disorders	0		0		0		0	0
		F64 Personality disorders	0		0		0		0	0
		F65 Personality disorders	0		0		0		0	0

Dimension		Psychopathology	Age							
		F66 Personality disorders	0		0		0		0	0
		F67 Personality disorders	0		0		0		0	0
		F68 Personality disorders	0		0		0		0	0
		F69 Personality disorders	0		0		0		0	0
	Sum for dimension 5		**0**		**0**		**0**		**0**	**0**
6	Psychotropic medication	N06B, N06A, N05B, N05A, N05CH	0		0		0		1	1
	Sum for dimension 6		**0**		**0**		**0**		**1**	**1**

Table S6: Comparisons of fit for one-to-six multi-group trajectory models for boys

Groups	BIC	BIC difference	Participants in each group (%)					
			1	2	3	4	5	6
1	-603783.1		100					
2	-575255.8	28527.3	11.2	88.8				
3	-568414.3	6841.5	71.6	22	6.4			
4	-566295.9	2118.4	54.6	35.1	5.5	4.8		
5	-562748.3	3547.6	59.9	3.8	27.9	5.3	3.1	
6	-561543.1	1205.2	23.4	9.9	54.9	5.2	3.5	3.0

BIC values closer to zero (as opposed to closer to negative infinity) are favoured in group-based trajectory modelling, to select the 'best' model. Trajectories were modelled using quadratic polynomials for all dimensions.
Estimation based on the pseudo population.
Note: All analyses were performed in SAS, where it is not possible to estimate a 7-group model using PROC TRAJ. However, we subsequently tested both 7 and 8 group models on a similar dataset with nearly the same sample and variables in STATA. These analyses indicated that additional groups resulted in very small group sizes and did not meaningfully improve the model.

Table S7: Comparisons of fit for one-to-six multi-group trajectory models for girls

Groups	BIC	BIC difference	Participants in each group (%)					
			1	2	3	4	5	6
1	-900999.2		100					
2	-866334.1	34665,1	85.9	14.1				
3	-857242.6	9091,5	67.1	25.7	7.2			
4	-851616.3	5626,3	7.3	58.9	30.3	3.5		
5	-849926.4	1689,9	44.3	17.1	27.6	6.7	4.3	
6	-846834.4	3092	34.1	16.6	35.0	6.3	5.8	2.2

BIC values closer to zero (as opposed to closer to negative infinity) are favoured in group-based trajectory modelling, to select the 'best' model. Trajectories were modelled using quadratic polynomials for all dimensions.
Estimation based on the pseudo population.
Note: All analyses were performed in SAS, where it is not possible to estimate a 7-group model using PROC TRAJ. However, we subsequently tested both 7 and 8 group models on a similar dataset with nearly the same sample and variables in STATA. These analyses indicated that additional groups resulted in very small group sizes and did not meaningfully improve the model.

Supplementary material

Table S8. Sociodemographic characteristics according to trajectory groups

		No Significant Problems	Minimal Difficulties	Neurotypical with Rising Behavioural Problems	Early Behavioural difficulties / Non-clinical Behavioural Difficulties	Adolescent-specific Psychopathology	Neurodiverse	Missing	P[b]
	N in groups:	12 923	24 101	5 582	3 018	2 274	1 463		
		% (n)	% (n)	% (n)	% (n)	% (n)	% (n)		
Maternal age at birth									
	16-19	0.1 (13)	0.3 (69)	0.5 (26)	1.1 (32)	0.4 (9)	1.3 (18)	76	<0.001
	20-24	4.5 (577)	7.1 (1 714)	9.3 (516)	12.1 (366)	8.3 (188)	13.7 (201)		
	25-29	36.5 (4 707)	39.2 (9 432)	38.6 (2 148)	41.1 (1 240)	36.8 (836)	39 (571)		
	30-34	41 (5 283)	38.3 (9 222)	37.5 (2 088)	32.6 (981)	38.1 (864)	32 (468)		
	35-39	16.9 (2 174)	14.2 (3 407)	13.3 (741)	12.4 (374)	15.3 (348)	13 (190)		
	40+	1.2 (149)	0.9 (222)	0.9 (51)	0.7 (21)	1.1 (24)	1 (15)		
Parity									
	0	35.9 (4 628)	36.8 (8 842)	32.9 (1 828)	38.9 (1 170)	35.6 (807)	36.3 (530)	148	<0.001
	1	32.4 (4 172)	34.5 (8 282)	35.1 (1 953)	34.1 (1 025)	34.7 (786)	32.9 (480)		
	≥2	31.7 (4 087)	28.7 (6 906)	32 (1 778)	27 (814)	29.7 (675)	30.8 (450)		
Maternal psychiatric diagnosis[a]									
	Not present	97.1 (12 544)	96.3 (23 205)	95.6 (5 337)	94.1 (2 839)	93.7 (2 131)	92.3 (1 350)	0	<0.001
	Present	2.9 (379)	3.7 (896)	4.4 (245)	5.9 (179)	6.3 (143)	7.7 (113)		
Paternal psychiatric diagnosis[a]									
	Not present	97.6 (12 614)	96.8 (23 326)	96.2 (5 367)	95.6 (2 886)	95.5 (2 171)	93.2 (1 363)	0	<0.001
	Present	2.4 (309)	3.2 (775)	3.9 (215)	4.4 (132)	4.5 (103)	6.8 (100)		
Marital status									

Supplementary material

		No Significant Problems	Minimal Difficulties	Neurotypical with Rising Behavioural Problems	Early Behavioural difficulties / Non-clinical Behavioural Difficulties	Adolescent-specific Psychopathology	Neurodiverse	Missing	P[b]
	Married/cohabiting	97.7 (12 626)	97.2 (23 419)	95.5 (5 331)	95 (2 867)	95.2 (2 164)	91.9 (1 345)	0	<0.001
	Not married/not cohabiting	2.3 (297)	2.8 (682)	4.5 (251)	5 (151)	4.8 (110)	8.1 (118)		
Maternal socioeconomic status[c]									
	Primary school	5.1 (660)	9.1 (2 190)	13.1 (731)	18.1 (541)	12.2 (277)	20.8 (304)	174	<0.001
	Vocational	39.4 (5 082)	47.1 (11 301)	50.1 (2 789)	52.8 (1 579)	44.8 (1 015)	49.8 (727)		
	Short-cycle	5.2 (668)	6.0 (1 449)	5.2 (289)	5 (150)	5.5 (124)	4.4 (64)		
	Medium cycle	36 (4 638)	28.4 (6 814)	23.8 (1 325)	17.9 (535)	28.8 (653)	18.9 (275)		
	Post graduate	13.7 (1 760)	9.07 (2 177)	7.5 (415)	6 (180)	8.5 (194)	6.1 (89)		
	Research/phd	0.7 (84)	0.4 (83)	0.3 (14)	0.2 (6)	0.2 (5)	0 (0)		
Alcohol consumption during pregnancy									
	Non-drinking	51.7 (6 606)	53 (12 633)	55 (3 025)	56.8 (1 697)	57.2 (1 287)	58.4 (841)	544	<0.001
	1-3 units/week	45.6 (5 831)	45 (10 640)	42.4 (2 334)	41.1 (1 227)	40.7 (917)	39.2 (564)		
	>3 units/week	2.7 (341)	2.5 (585)	2.6 (143)	2.1 (63)	2.1 (48)	2.4 (35)		
Child, small for gestational age									
	No	97.9 (12 438)	97 (23 027)	96.9 (5 313)	95.6 (2 843)	96.7 (2160)	94.7 (1 367)	770	<0.001
	Yes	2.1 (272)	3 (720)	3.1 (171)	4.4 (130)	3.3 (74)	5.3 (76)		
Symptoms of colic									
	No	93.8 (9 697)	92.7 (17 833)	92.1 (4 137)	90.5 (2 149)	90.4 (1 658)	88.5 (1 032)	9 919	<0.001
	Yes	6.2 (636)	7.3 (1 406)	7.9 (356)	9.5 (227)	9.6 (177)	11.5 (134)		

[a] Psychiatric diagnosis before birth of DNBC child

[b] Chi-square test

[c] Education in Denmark ranges from primary school (basic compulsory education), through vocational education and training programmes, short-cycle higher education, medium-cycle higher education (bachelor's degrees), post graduate studies (master's degrees), to research and PhD-level doctoral degrees.

Supplementary material

Table S9: GRoLTS-checklist

	Checklist Item	Reported?
1.	Is the metric of time used in the statistical model reported?	Yes. An example is presented in Table S5
2.	Is information presented about the mean and variance of time within a wave?	No
3a.	Is the missing data mechanism reported?	Yes. The specifics of this information are explained in the method section.
3b.	Is a description provided of what variables are related to attrition/missing data?	Yes. The specifics of this information are explained in the method section.
3c.	Is a description provided of how missing data in the analyses were dealt with?	Yes. The specifics of this information are explained in the method section.
4.	Is information about the distribution of the observed variables included?	Yes, Table S2
5.	Is the software mentioned?	Yes. We used SAS' Proc Traj
6a.	Are alternative specifications of within-class heterogeneity considered (e.g., LGCA vs. LGMM) and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	We applied latent class growth modelling. We visually inspected the trajectories with one up to six groups to determine the best-fitting model. The paper has provided a thorough explanation of the procedure's details.
6b.	Are alternative specifications of the between-class differences in variance–covariance matrix structure considered and clearly documented? If not, was sufficient justification provided as to eliminate certain specifications from consideration?	
7.	Are alternative shape/functional forms of the trajectories described?	No. The paper does not report alternative shapes/functional forms, but we have tested alternative specifications and examined the models.
8.	If covariates have been used, can analyses still be replicated?	Our model was an unconditional model.
9.	Is information reported about the number of random start values and final iterations included?	No.
10.	Are the model comparison (and selection) tools described from a statistical perspective?	Yes, Table 1
11.	Are the total number of fitted models reported, including a one-class solution?	Yes, Table S6 & S7
12.	Are the number of cases per class reported for each model (absolute sample size, or proportion)?	Yes, Table S6 & S7
13.	If classification of cases in a trajectory is the goal, is entropy reported?	No. Classification of cases in a trajectory was not the goal
14a.	Is a plot included with the estimated mean trajectories of the final solution?	Yes, Figure 2 and Figure 3
14b.	Are plots included with the estimated mean trajectories for each model?	No
14c.	Is a plot included of the combination of estimated means of the final model and the observed individual trajectories split out for each latent class?	No
15.	Are characteristics of the final class solution numerically described (i.e., means, SD/SE, n, CI, etc.)?	The percentages are shown in Table 2 and Table S8.
16.	Are the syntax files available (either in the appendix, supplementary materials, or from the authors)?	No. However, there is a detailed description of it in the method section and supplementary materials. The code utilized in this study corresponds to the code provided in the proc traj statement. It has been sourced from the official public repository, accessible

		at https://www.andrew.cmu.edu/user/bjones/index.htm, ensuring fidelity to the original methodology and implementation. Figures were created in R.

Rens van de Schoot, Marit Sijbrandij, Sonja D. Winter, Sarah Depaoli & Jeroen K. Vermunt (2017) The GRoLTS-Checklist: Guidelines for Reporting on Latent Trajectory Studies, Structural Equation Modeling: A Multidisciplinary Journal, 24:3, 451-467, DOI: 10.1080/10705511.2016.1247646

Table S9: STROBE Statement

	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1-2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4
Methods			
Study design	4	Present key elements of study design early in the paper	5
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up (*b*) For matched studies, give matching criteria and number of exposed and unexposed	5
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	6-7
Bias	9	Describe any efforts to address potential sources of bias	8
Study size	10	Explain how the study size was arrived at	6
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	7-8
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding (*b*) Describe any methods used to examine subgroups and interactions (*c*) Explain how missing data were addressed (*d*) If applicable, explain how loss to follow-up was addressed (*e*) Describe any sensitivity analyses	7
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed (b) Give reasons for non-participation at each stage (c) Consider use of a flow diagram	5 and Figure 1
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders (b) Indicate number of participants with missing data for each variable of interest (c) Summarise follow-up time (eg, average and total amount)	Supplementary material

Outcome data	15*	Report numbers of outcome events or summary measures over time	Supplementary material
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included (*b*) Report category boundaries when continuous variables were categorized (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	8, Table 1 and 2
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	
Discussion			
Key results	18	Summarise key results with reference to study objectives	9
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	17
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	18
Generalisability	21	Discuss the generalisability (external validity) of the study results	17-18
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Under the section "Funding"

*Give information separately for exposed and unexposed groups.

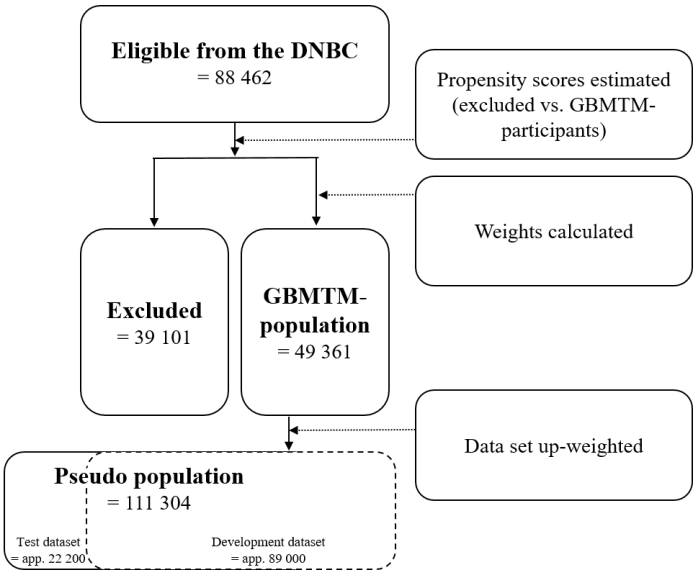
Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

Note 1: Weighted analyses for the GBMT model analyses

We applied inverse probability weights in our GBMT model analyses. The weights were constructed and applied as described in the following:

1. *Propensity scores*: We used logistic regression to estimate probability of having sufficient data (yes vs. no) to be included in the group-based multi-trajectory analyses, i.e., being included in the study population (=49 361) out of the eligible live-born term singletons without congenital anomalies, conditions affecting CNS and global developmental delay from Danish National Birth Cohort sample of 88 462 children (see flowchart for exclusion steps). The following predictors were included in the logistic regression: maternal psychiatric diagnosis (yes/no(ref.)), paternal psychiatric diagnosis (yes/no(ref.)), child psychiatric diagnoses (yes/no(ref.)), parity (0(ref.), 1, ≥2), maternal age (16-19, 20-24, 25-29 (ref.), 30-34, 35-39, 40+), alcohol consumption during pregnancy (non-drinking (ref.), 1-3 units/week, >3 units/week), maternal smoking during pregnancy (non-smoking (ref.), smoking quitter, smoker), maternal educational level (primary school (ref.), vocational Short-cycle, medium cycle, postgraduate/Research/PhD.), urban/rural residence (capital (ref.), capital suburb, provincial city (>100 000 inhabitants), provincial city (>10 000 inhabitants), rural area), and quantiles of household income (1(lowest)(ref.)-5(highest), number of answered pregnancy interviews (1-4 (ref.)). The regressions were gender stratified.
2. *Weights*: Weights were calculated for the included individuals based on their estimated propensity score (weight=1/propensity score).
3. *Cloning*: The GBMT analyses were fitted on cloned data where individuals with a low probability of having all required data were up-weighted, e.g., an individual with a weight of three was reproduced three times, thus appearing three times in the dataset. This data cloning resulted in a pseudo-population of 111 304 compared to the 49 361 unique individuals with complete data on the minimum number of variables needed to run the GBMT analyses.

Supplementary material



Note: The pseudo-population (=111 304) became slightly larger than the original data set with eligible DNBC participants (=88 462) because the weighting procedure included only the rounded weight values. Since the GBMTM analyses were developed on 80% of the pseudo-population, the data analysed in the development phase was similar to the original sample size. We did not apply down-sampling before running the final analyses. The impact on the CIs and p-values was minimal when transferring the model developed on the 80% to the inflated complete (pseudo) population.