

Change in the severity and impact of child psychiatric disorders over 3 years: an analysis of trends using the national child mental health surveys of England from 1999, 2004, and 2017

Jessica M. Armitage,^{1,2,3†} Tamsin Newlove-Delgado,^{3†} Tamsin J. Ford,⁴ Sally McManus,^{5,6} Sharon A. S. Neufeld,⁴ Yuanfei Liu,⁴ and Stephan Collishaw^{1,2}

¹Wolfson Centre for Young People's Mental Health, Cardiff University, Cardiff, UK; ²Division of Psychological Medicine and Clinical Neurosciences, MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University School of Medicine, Cardiff University, Cardiff, UK; ³Medical School, University of Exeter, Exeter, UK; ⁴Department of Psychiatry, Cambridge University, Cambridge, UK; ⁵National Centre for Social Research, London, UK; ⁶School of Health and Medicine, City St George's, University of London, London, UK

Background: Child and adolescent psychiatric disorders are common, now affecting at least one in eight children. Evidence points towards an increase in prevalence over time, but very few studies have tested whether longitudinal outcomes for children with mental health problems have changed. This study is the first to test whether mental health and functional outcomes for children with a diagnosed psychiatric disorder have improved or worsened across time in England. **Methods:** Three national probability surveys were undertaken in England in 1999, 2004, and 2017, each with respective 3-year follow-ups in 2002, 2007, and 2020. We include children at baseline aged 5–13 years and compare outcomes for those with baseline psychiatric diagnosis (mental health disorder or ADHD). Three-year mental health symptom severity and functional impairment were assessed using the parent- and self-rated total difficulties score and impact supplement from the Strengths and Difficulties Questionnaire. **Results:** Across all surveys, having a psychiatric disorder at baseline predicted greater difficulty and impact scores at follow-up. The magnitude of these associations increased sequentially over time, such that having a disorder at baseline was more strongly associated with parent- and self-rated difficulties in the 2017 survey compared with the 1999 and 2004 surveys. Children in the 2017 survey were also more likely to report a greater impact of problems. Patterns of individual change also showed a temporal gradient. For children with a psychiatric disorder in 1999, 34% experienced deterioration in mental health at the three-year follow-up, compared to 52% of those with a disorder in 2017. The proportion of those who improved decreased correspondingly. **Conclusions:** Children with a psychiatric condition are experiencing increased difficulties in their daily lives compared to previous generations. Understanding why children are now more likely to experience impaired functioning is a major public concern. **Keywords:** Adolescence; developmental epidemiology; public health.

Background

Child and adolescent psychiatric disorders are common and increasing. In 2020, NHS Digital reported that the number of 5- to 16-year-olds in England living with a probable mental health disorder rose from one in eight in 2017 to one in six in 2020 (Vizard et al., 2020). This increase is especially concerning as psychiatric conditions in children and young people are associated with an increased risk for later mental health difficulties, impaired social relationships, educational underachievement, and physical ill health (Copeland, Alai, Jonsson, & Shanahan, 2021). Yet despite the host of epidemiological studies pointing towards generational shifts in mental health, notably an increase in depression and anxiety over time (Armitage et al., 2023; Sadler et al., 2018), only very few have tested whether the longitudinal outcomes for

children who have a mental health disorder have improved or worsened.

Research comparing outcomes for different generations of children with mental health difficulties based on equivalent screening thresholds in the UK has suggested that outcomes are more negative among individuals born more recently (Sellers et al., 2019). In Sellers et al.'s study, child mental health difficulties were more strongly associated with poorer social functioning, lower academic attainment, and increased persistence of mental health problems in adolescence among more recent cohorts Sellers et al. (2019). Similar results have been found when comparing outcomes in adulthood (Thompson, Richards, Ploubidis, Fonagy, & Patalay, 2023). However, no epidemiological study has yet used a clinically-validated diagnostic instrument to assess whether difficulties and impact have changed for those meeting criteria for psychiatric disorder, nor assessed whether any such patterns of change vary for different population subgroups, or for different psychiatric conditions.

†Joint first authorship.

Conflict of interest statement: No conflicts declared.

Testing how outcomes have changed for those with a disorder, and who is most affected will provide insight into the effectiveness of our population mental health systems. It will also inform planning of future services (Collishaw & Sellers, 2020). In the context of debates around under- or over-diagnosis of problems in children and the role of rising awareness, investigating whether the prognosis of having a psychiatric disorder has worsened for children with psychiatric disorders therefore represents an important public health priority.

This study aims to understand the severity of future mental health difficulties and functional impairment faced by children with a psychiatric disorder, and how these longitudinal outcomes have changed across time. A second aim was to explore how changes in three-year outcomes vary by age, gender, ethnicity, and baseline disorder type. To do this, we use unique data from three nationally representative surveys undertaken in 1999, 2004, and 2017 in England, each with three-year follow-up surveys in 2002, 2007, and 2020.

Methods

Sampling strategy

National probability samples of children in England were analyzed from three cohorts that were part of the same survey series—the British Child and Adolescent Mental Health Surveys (BCAMHS) in 1999 and 2004, and the NHS England Mental Health of Children and Young People in England (MHCYP) 2017 (Ford et al., 2020). All included three-year longitudinal follow-ups, which we use to examine individual outcomes. The 3-year follow-ups sampled all children included at baseline, except for the BCAMHS 2002 follow-up, which included all children with a disorder at baseline and a randomly selected third of those without baseline disorder (Green, McGinnity, Meltzer, Ford, & Goodman, 2005).

The 1999 and 2004 BCAMHS samples were selected from the Child Benefit Register, which had universal coverage at that time. The final samples represented approximately 90% of all British children (Ford et al., 2020). The MHCYP 2017 sample was drawn from the National Health Service (NHS) Patient Register as Child Benefit was no longer universal. Children in all surveys were selected from a stratified probability sample of postcode sectors (Ford et al., 2020). All three surveys included weights to account for sample design and non-response at baseline, as well as weighting to account for differential loss to follow-up based on factors associated with non-response including region, age, sex, and disorder (Meltzer, Gatward, Corbin, Goodman, & Ford, 2003; Parry-Langdon, 2008). Further information on response is presented in Appendix S1.

Data from all baseline surveys were collected from parents, children (aged ≥ 11 years), and teachers (the family nominated the teacher who knew the child best). Teachers were sent questionnaires. Parents and children were interviewed face-to-face at home by trained lay interviewers.

The 2002 and 2007 follow-ups used the same procedure; however, the 2020 follow-up to the 2017 survey was conducted online due to the pandemic, with fieldwork taking place in July and August 2020.

The current study included children aged 5 to 13 years at baseline who had information on disorder status and three-year follow-up data available (BCAMHS 2002, $n = 1,824$; BCAMHS 2007, $n = 3,657$ and MHCYP 2020,

$n = 2,053$). We restricted our sample to children living in England, as MHCYP only included England.

All secondary analyses were approved by the University of Exeter Medical School REC (Nov20/D/270). MHCYP2017 and 2020 data were used under a data sharing agreement with NHS England (DARS-NIC-424336-T7K7T-).

Measures

International Classification of Diseases (ICD-10) psychiatric disorders. Psychiatric disorders were assigned across all three samples at baseline by trained clinical raters, using the data from parents, children and teachers gathered using the Development and Well-Being Assessment (DAWBA), a structured diagnostic assessment based on ICD and Diagnostic Statistical Manual (DSM) criteria (Goodman, Ford, Richards, Gatward, & Meltzer, 2000). We include both psychiatric disorders and ADHD under the shorthand “umbrella term” of psychiatric disorder, in line with the baseline surveys, which used standardized diagnostic criteria from the ICD-10. Further information about the procedures for clinical rating of the DAWBA can be found elsewhere (Ford et al., 2020). Within the present study, we included individuals meeting the criteria for at least one ICD-10 psychiatric disorder at baseline, using a harmonized list of disorders included in the DAWBA in 1999, 2004, and 2017 (see Armitage, Collishaw, & Sellers, 2024 for further details). This included: anxiety disorders, depressive disorders, conduct disorders (including Oppositional Defiant Disorder (ODD)), hyperkinetic disorders/ADHD, eating disorders, and tic disorders. A full list of DAWBA topic coverage of disorders across the three years is available in the 2017 Survey Methods report (Vizard et al., 2020). We also compared outcomes for children meeting the criteria for either an emotional disorder (anxiety or depressive disorder), a behavioral disorder (conduct disorder or oppositional defiant disorder (ODD)), a hyperkinetic disorder (attention deficit and hyperactivity disorder (ADHD)), or comorbid disorders (those meeting diagnostic criteria from two or more of these diagnostic groupings).

Difficulties and impact. Difficulties and functional impacts experienced by the child at the 3-year follow-up were assessed using the Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997). This 25-item scale splits into five subscales: emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior. All subscales (except prosocial) are summed to form a total difficulty score ranging from 0 to 40, with higher scores indicating greater difficulties. We used total difficulty scores reported by the parent (when child was 8–16 years) and the child (for 11- to 16-year-olds). Teacher ratings were not available across all surveys and therefore excluded. We derived “change in difficulties” capturing change in mean parent SDQ total score from baseline to follow-up, following the same procedure used in the MHCYP surveys (Newlove-Delgado et al., 2021). Deterioration was the term used where there was an increase of three or more in the parent-rated SDQ score, while a decrease of three or more indicated improvement. If the SDQ score was no more than two points higher or lower than baseline, the child was classified as having experienced “stable” mental health scores.

The SDQ impact supplement comprises five items, each rated 0—not at all or only a little, 1—a medium amount, or 2—a great deal. These capture overall distress and impairment experienced by the young person, as well as the degree to which symptoms interfere with four aspects of life (friendships, classroom learning, home life, and leisure activities). The items are summed to generate an overall impact score ranging from 0 to 10, with higher scores indicating greater impact. As with the SDQ total difficulty scale, we used reports from both parents and children.

Choice of primary outcome measure

The SDQ has been validated as a tool for both cross-cohort comparisons and longitudinal research (Armitage et al., 2023; Rigin et al., 2024). This sample had SDQ scores available at baseline and three-year follow-up, allowing us to assess deterioration or improvement in problems.

Sociodemographic characteristics

Sociodemographic characteristics included gender, age, ethnicity. We also compared baseline characteristics of weighted and unweighted samples using measures related to socioeconomic status, including family type (lone parent or none), housing tenure (whether families owned or rented their accommodation), and household income. Household income was assessed using a showcard and included earnings from employment and benefits. Analyses then categorized household income as low (bottom quintile) or medium-high (above bottom quintile).

Statistical analyses

Measurement invariance. Establishing measurement invariance supports the accuracy of comparing associations or means across groups or time by showing there are no systematic differences in a measure's factor structure, item loadings, or item endorsements across groups or time. We have previously shown parent- and self-reported SDQ to be invariant by sex, age, and rater in 2017 MHCYP (master's thesis, pre-print available shortly). For the present paper, we tested measurement invariance in the SDQ across the three main surveys (1999, 2004, 2017) in both parents and adolescents, and also from 2017 to 2020 follow-up (statistical methods in Appendix S1).

Descriptives. We first compared the sociodemographic profiles of those included at both baseline and follow-up across the three surveys.

Main analyses. Linear regressions were used to examine associations between any disorder at baseline and SDQ score at follow-up, with tests of interactions by cohort used to test change in the strength of these associations across time. Analyses were carried out separately using parent-rated SDQ total difficulty scores and then self-reported difficulty scores. Further analyses then compared associations between any disorder at baseline and parent- and self-rated SDQ impact scores at follow-up.

Main analyses were repeated after stratification by age, gender, and ethnicity, and three-way interactions between cohort, disorder, and each demographic variable were used to test if any change in the strength of associations between disorder and outcome over time differed by demographic factor.

Individual-level change in parent-rated SDQ scores from baseline to follow-up was compared between the three surveys. These were conducted by comparison of proportions tests and were restricted to those with any disorder at baseline.

Sensitivity analyses. Analyses investigated change in associations between presence of specific disorders at baseline and SDQ total and impact scores at follow-up.

Adjusting for survey design and probability weights. All analyses were conducted in Stata version 17 (StataCorp, 2021). Sampling weights as provided with the datasets were used in all main analyses to adjust for differences in the sampling frame and differential follow-up rates by region and strata (please see [sampling strategy](#) above for more

information). Table 1 shows the sociodemographic characteristics at baseline of those with complete and missing data at follow-up across the three surveys and Table S1 displays an unweighted comparison of sociodemographic characteristics. Table S2 displays the relationships between mental and physical health, socioenvironmental, and sociodemographic variables in the 2017 survey and missingness in SDQ in 2020.

Role of the funding source

The funders of the study had no role in this study.

Results

Measurement Invariance

The highest level of measurement invariance (strict, or residual level) was established for both parent and self-reported SDQ scores across the 3 main surveys (Table S3). For the follow-up period from 2017 to 2020, residual measurement invariance was also demonstrated in parent-reported SDQ, and partial scalar (strong) measurement invariance in self-report SDQ supported minimal bias in the total difficulties score (Table S4 and Appendix S2).

Profile and representativeness of the three cohorts at baseline and follow-up

The sociodemographic profiles of those included at both baseline and follow-up across the three surveys are shown in Table S1.

Comparison of risk associated with disorder by cohort

Compared to those without a disorder at baseline, having any psychiatric disorder was associated with increased parent- and self-rated SDQ total difficulty score at follow-up across all surveys, as well as higher SDQ impact scores at follow-up (Table 2). Cohort by child mental health interactions indicated that this risk was significantly greater for those with a disorder in 2017 compared to 1999 or 2004, with the difference between those with and without a disorder on average 2.48 points (SE = 0.87) greater on the parent-rated SDQ at follow-up for the 2017 compared to those first assessed in 1999 (see Table 2 and Table S5 for unweighted results).

Analyses stratified by gender revealed some differences. While SDQ scores were consistently higher among those with disorder compared to those without, comparisons across time showed that the association of baseline disorder with parent-reported total difficulty scores at follow-up was greater in 2017 compared to 1999 among males but not females, although the pattern of results was the same for females (see Figure 1). For females, the association of baseline disorder with total difficulty scores at follow-up was greater in 2017 compared to 1999 when self-reported (see Figure 1 and Table S6).

	BCAMH 1999 & 2002			BCAMH 2004 & 2007			MHCYP 2017 & 2020		
	Complete (n = 1,824)	Complete weighted	Missing/not sampled (n = 5,517) ^a	Complete (n = 3,657)	Complete weighted	Missing (n = 1,690)	Complete (n = 2,053)	Complete weighted	Missing (n = 3,159)
Age (mean)	9.0	8.9	8.9	9.1	9.2	9.2	8.8	8.8	8.9
% Female	47.2%	48.8%	51.1%	47.8%	47.5%	48.9%	50.6%	49.4%	48.7%
% From ethnic minority background	6.9%	7.9%	10.7%	11.4%	12.1%	22.0%	16.4%	19.3%	24.0%
% In rented accommodation	26.8%	24.2%	33.4%	24.8%	27.4%	41.8%	29.6%	30.9%	48.4%
% Lowest income quintile	21.6%	19.4%	24.3%	16.7%	19.2%	29.9%	17.4%	18.1%	28.1%
% With any psychiatric disorder	20.4% ^a	8.6%	4.0%	6.8%	8.3%	10.4%	7.3%	7.1%	10.2%
Mean parental total difficulty SDQ score	9.4 ^a	8.5	8.3	7.8	8.2	9.1	7.5	7.6	8.3

Sampling for the 3-year follow-up in the 1999 survey included all those with disorder at baseline and a third of those without—see [Methods](#).

	BCAMH 1999 & 2002 (<i>n</i> = 1,824)			BCAMH 2004 & 2007 (<i>n</i> = 3,657)			MHCYP 2017 & 2020 (<i>n</i> = 2,053)			Interaction by survey year			
	Disorder Mean (SD)	No disorder Mean (SD)	Difference Coefficient (SE)	Disorder Mean (SD)	No disorder Mean (SD)	Difference Coefficient (SE)	Disorder Mean (SD)	No disorder Mean (SD)	Difference Coefficient (SE)	1999 vs. 2004 Coefficient (SE)	1999 vs. 2017 Coefficient (SE)	2004 vs. 2017 Coefficient (SE)	
SDQ at follow-up	15.13 (7.78)	7.04 (5.14)	8.09*** (0.43)	16.88 (7.90)	7.39 (5.32)	9.49*** (0.52)	19.14 (7.79)	8.57 (5.93)	10.58*** (0.76)	1.40* (0.67)	2.48** (0.87)	1.09 (0.93)	
Parent-rated total difficulty scores	13.98	8.87	5.11***	15.08	9.56	5.52***	17.77	10.35	7.42***	0.41 (0.62)	2.31*	1.90*	
Self-rated total difficulty scores (11–16 only)	(5.72)	(4.90)	(0.41)	(5.50)	(4.97)	(0.47)	(6.19)	(5.69)	(0.81)		(0.91)	(0.94)	
Parent rated impact scores	2.05 (2.54)	0.33 (1.01)	1.72*** (0.14)	2.32 (2.73)	0.31 (0.97)	2.01*** (0.18)	3.76 (3.10)	0.64 (1.53)	3.12*** (0.29)	0.29 (0.23)	1.41*** (0.32)	1.12*** (0.34)	
Self-rated impact (11–16 only)	0.75 (1.51)	0.22 (0.80)	0.53*** (0.10)	0.98 (1.50)	0.22 (0.76)	0.76*** (0.13)	2.67 (2.77)	0.57 (1.31)	2.10*** (0.36)	0.22 (0.17)	1.56*** (0.38)	1.34*** (0.39)	

* $p < 0.05$
** $p < 0.01$
*** $p < 0.001$

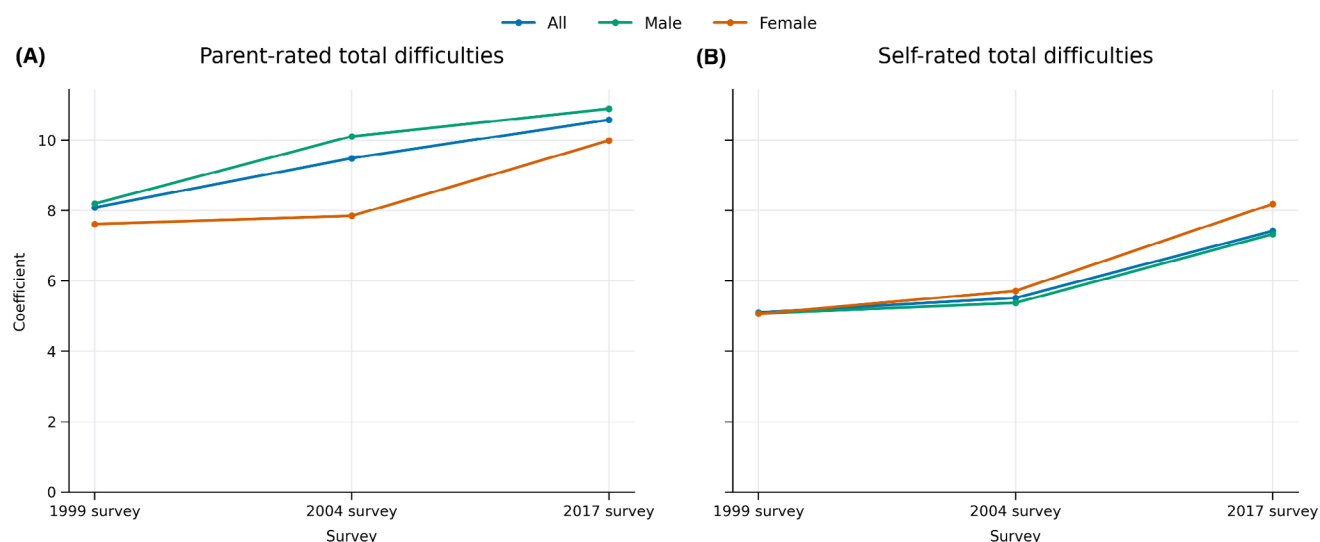


Figure 1 Comparison of regression coefficients for associations between any psychiatric disorder at baseline and mental health difficulties at follow-up. Note that parent-reports are based on 8–16 year olds at follow-up while self-reports are based on those aged 11–16 years. Differences between female self-rated scores in 1999 and 2017 were significant

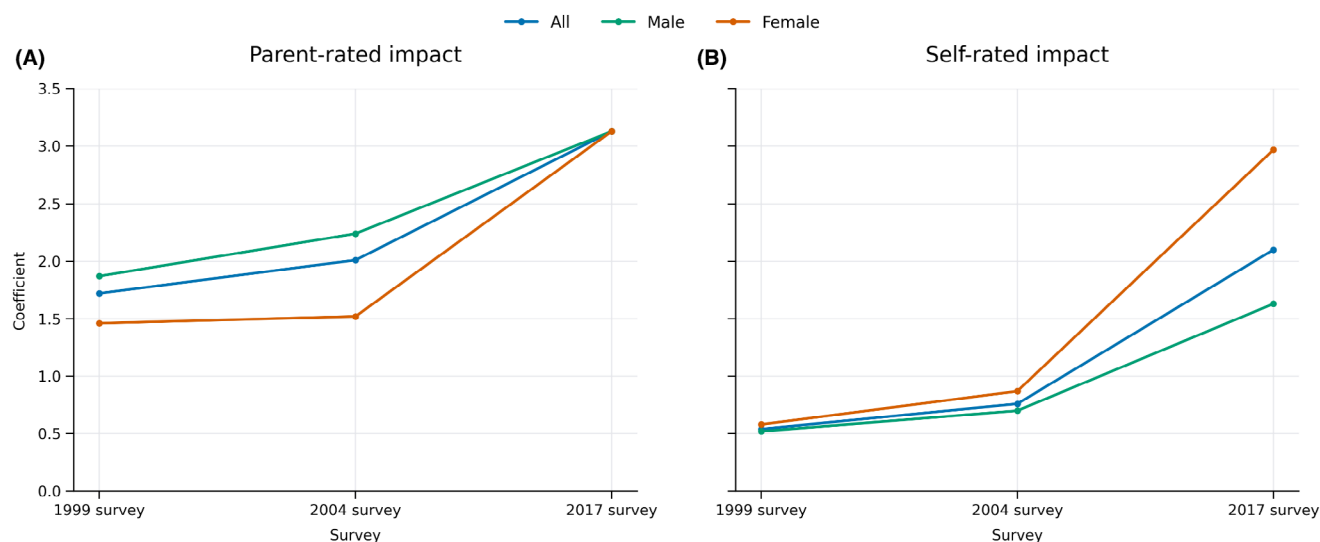


Figure 2 Comparison of regression coefficients for associations between any psychiatric disorder at baseline and mental health impact at follow-up. Note that parent reports are based on 8–16 year olds at follow-up while self-reports are based on those aged 11–16 years

Three-way interaction analyses found that gender differences in associations between baseline disorder and total difficulty scores over time were not significantly different (see Table S7). For males and females, baseline disorder was more strongly associated with impact scores at follow-up in 2017 compared to 1999 and 2004 for both parent- and self-reported (Figure 2).

Analyses stratified by age indicated increased associations between baseline disorder and difficulties at follow-up in more recent surveys both for younger (8-to-10 years) and older (11–16 years) children (see Table S8 and Figure S1). Three-way interactions suggested there were no systematic age-related differences in the main findings (Table S9).

When stratified by ethnicity, differences across surveys in associations between disorder and total

difficulty scores were only apparent among white young people (Table S10). Difference between self-reported difficulties among those with or without disorder decreased over time for non-white groups (Table S10).

Ethnic differences across time were shown to be significantly different within the interaction analyses (Table S11). Samples were small; therefore, results should be interpreted with some caution.

Comparison of individual change for those with disorder, by cohort

Changes in mental health difficulties between baseline and follow-up were compared among those with disorder at baseline using parent-rated SDQ total difficulty scores. Comparison across the three

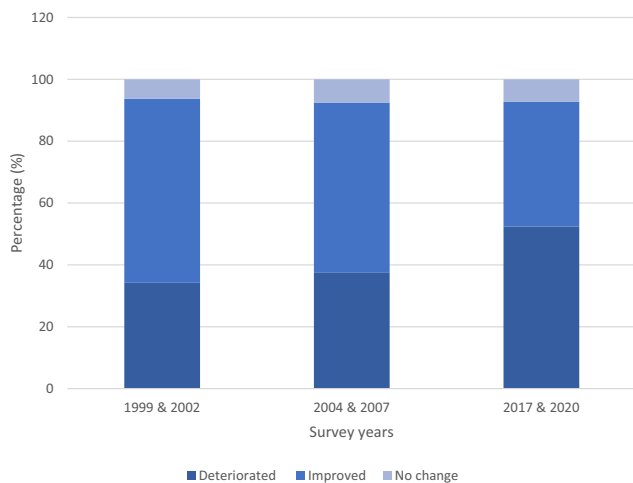


Figure 3 Comparison of three-year outcomes, as assessed using change in parent-rated Strengths and Difficulties Questionnaire (SDQ) total difficulty score across the three surveys and their follow-ups

surveys revealed an increase over time in the percentage of those whose mental health deteriorated from baseline to follow-up (see Figure 3). Compared to the two earlier surveys (1999 and 2004), in which 34.2% and 37.6% deteriorated at three-year follow-up, the latest survey (2017) had the highest proportion of adolescents who deteriorated between baseline and follow-up (52.4%), and the lowest proportion who improved (Table S12). Similar patterns were found when stratified by gender and age, and when analyses were restricted to individuals of white ethnicity (see Table S13). The very low number of children from other ethnic backgrounds meant we were unable to perform meaningful analyses for these groups (see—Limitations).

Sensitivity analyses, other disorders

Sensitivity analyses stratified by disorder found a similar pattern of increasing difficulty and impact scores over time, but with wide confidence intervals. There were no significant differences across surveys in association between disorder and total difficulty scores at follow-up for those with an emotional, behavioral, or hyperkinetic disorder, but there was an increased association between 1999 and 2017 with parent-rated impact scores for those with an emotional disorder (see Tables S14–S17). Individuals with comorbid disorders had greater parent-rated total difficulty scores at follow-up in 2017 compared to 1999 (see Table S17). It is important to note that these subgroup comparisons are likely underpowered due to small sample sizes when considering individual disorder categories.

Discussion

This paper presents a novel analysis of change over the first twenty years of the 21st Century in difficulties and functional impact of disorder among

children with psychiatric disorders in a series of national probability surveys. Our headline finding is that children with a disorder at baseline in the latest (2017) survey were more likely to experience deterioration in their mental health in the three years from baseline to follow-up than children in earlier surveys. This applied both to total difficulty scores on the SDQ and to impact scores. Our findings also demonstrate that experiencing a psychiatric disorder became more strongly associated with problematic mental health and functional outcomes in the most recent generation.

Importantly, these findings are bolstered by evidence of measurement invariance in SDQ across the three surveys and from 2017 to 2020 follow-up. This supports the suggestion that difficulties are being reported in a similar way even during the COVID-19 pandemic and are largely unbiased by measurement artifact, as also shown for psychological distress in adults in the UK Longitudinal Household Survey (Schlechter, Ford, McManus, & Neufeld, 2023).

Our study is the only one we are aware of which examines change in difficulties and functional impact for children in a population sample with data on disorder status at baseline, using standardized diagnostic assessments. This complicates comparison with other literature, which uses screening thresholds at baseline. However, these studies have also reported poorer mental health, social and educational outcomes amongst children born more recently (Sellers et al., 2019). Our findings also accord with our linked study which used the same baseline surveys and found higher concurrent functional impact and total difficulty scores in children with a disorder in 2017, compared to children in 1999 and 2004 (Armitage, Newlove-Delgado, Ford, McManus, & Collishaw, 2025).

If our findings represent a “true” worsening in outcomes, why might this be? By coincidence, the 3-year follow-up for the 2017 MHCYP survey was in the planning stages at the beginning of the COVID-19 pandemic. Fieldwork took place in late July and August of 2020, shortly following easing of lockdown measures. This 2020 follow-up clearly demonstrated a sharp increase in the prevalence of probable mental disorder at population level between 2017 and 2020. Further follow-ups of the sample in 2021, 2022, and 2023 showed that this increase was maintained across the immediate post-pandemic period (Newlove-Delgado et al., 2021). On balance, it appears likely that the pandemic contributed to the secular trend we observe of poorer outcomes in the most recent survey, but other explanatory factors may also play a role.

It is noteworthy that secular changes in concurrent functional impacts and mental health problem severity associated with disorder were observed in these same cohorts prior to the pandemic (Armitage et al., 2025), and other evidence indicates longer-term deterioration in psychosocial outcomes

for children with mental health difficulties evident over the second half of the 20th century (Sellers et al., 2019). Evidence points toward co-occurring shifts in parental emotional problems, eating disorders and weight-control behaviors in young people, as well as increases in school-related stress and family poverty (Armitage et al., 2024). Other evidence highlights growing inequalities in children's health and mental health (Collishaw, Furzer, Thapar, & Sellers, 2019). The economic crisis of 2008 and subsequent periods of public spending cuts are likely to have disproportionately impacted health programs and mental health and social services used by children from the most disadvantaged families (Children's Commissioner, 2021; Stuckler, Reeves, Loopstra, Karanikolos, & McKee, 2017). It is worth noting that only a third of children with a disorder from low-income backgrounds experienced improvement or stability in their symptoms between 2017 and 2020, a substantially lower proportion than for previous surveys or for middle and high income children with disorder in 2017 (where around half did). Whilst outside the scope of this paper, it is therefore also relevant to consider whether changes and inequalities in access to early intervention and/or treatment for mental health problems could also have resulted in poorer outcomes for our most recent cohort (Children's Commissioner, 2021).

Methodological limitations or differences may have influenced our findings, with many of these relating to the nature of the 2020 MHCYP survey, which was constrained by COVID-19. Firstly, we were unable to examine 3-year outcomes in terms of persistence or remission of disorder, as the MHCYP 2020 follow-up did not include the DAWBA, only the SDQ. Secondly, the timing of the MHCYP 2020 follow-up meant that teachers were not invited to participate; hence, we could not compare trends in teacher-rated outcomes over time, which would have provided additional triangulation. Thirdly, MHCYP 2020 was an online survey compared to the previous surveys conducted face to face. However, research suggests there is not a significant mode effect on completion of our main outcome measure, the SDQ, when comparing online to in-person completion (Mauz, Houben, Hölling, & Schlack, 2014).

Next, the lack of "keeping in touch" between 2017 and 2020 resulted in marked attrition between 2017 and 2020, a response rate of 46% in 2020, compared with 73% in BCAMHS 2007. Whilst weighting was applied to adjust for some measured demographic factors associated with loss to follow-up (such as region and housing tenure), this does not account for unmeasured factors associated with survey response and is likely to fail in restoring the sample to be fully representative in nature. Further, loss to follow-up in longitudinal surveys is typically greater for socially disadvantaged families and for children with a greater burden of mental health, health and educational difficulties (Robson et al., 2025), and

therefore, our analyses may be conservative, underestimating deterioration in outcomes for the most recent cohort.

A further limitation relates to the smaller sample sizes for examination of outcomes by individual disorders and for children of ethnic minority. Our study provides suggestive evidence of differential patterns of trends in outcomes for different groups of young people with disorders, but further better-powered research is needed to confirm how these patterns vary by gender, ethnicity, socio-economic background, and disorder type.

A key implication for future research is the importance of data linkage for future cohorts to enable examination of longer-term and wider outcomes, including linkage to outcomes such as self-harm, suicide, and severe mental illness, as well as to broader health, education, and employment data. Such linkage may also help to fill in the gaps for those that are otherwise lost to follow-up. Future national cohort studies should invest in meaningful engagement with the children, young people and families in their sample, through "keeping in touch" contacts, co-production, and incentivization, to address increasing rates of non-response and attrition. Future surveys should also focus on recruitment and retention of sufficient numbers of participants from different ethnic and cultural backgrounds to enable fine-grained analysis, or otherwise, ensure suitable "boosters" or complementary studies are commissioned.

Despite these limitations, we stress the value of our study which is the first to compare outcomes for children with disorders across three national cohorts. For children, three years represents both a substantial proportion of their young lives and encompasses significant periods of developmental change. Our finding that children with mental health conditions now appear to experience poorer mental health and functional outcomes than previously therefore has concerning implications in terms of the effectiveness of our population mental health systems, serving as a wakeup call. In line with other international research, all three surveys demonstrate that poor outcomes are not inevitable, and some children with disorders will show improvement over time, but a significant (and growing) proportion now experience deterioration in their mental health and psychosocial functioning. Our findings suggest that researchers and services must concentrate on addressing the factors that contribute to development, maintenance, and exacerbation of problems, and strengthening the system from prevention to provision of effective treatment.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article:

Appendix S1. Supplementary methods.

Appendix S2. Supplementary results.

Table S1. Comparison of sociodemographic characteristics across the three surveys (unweighted).

Table S2. Relationships^a between mental and physical health, socioenvironmental, and sociodemographic variables in the 2017 survey and missingness in SDQ in 2020.

Table S3. Multigroup measurement invariance testing of 3 main surveys (1999, 2004, 2017).

Table S4. Longitudinal measurement invariance testing, 2017 survey to 2020 follow-up.

Table S5. Comparison of associations between any disorder and outcomes across the three surveys (unweighted).

Table S6. Comparison of associations between any disorder and outcomes across the three surveys, by gender.

Table S7. Comparison of gender differences over time in associations between any disorder and outcomes across the three surveys.

Table S8. Comparison of associations between any disorder and outcomes across the three surveys, by age.

Table S9. Comparison of age differences over time in associations between any disorder and outcomes across the three surveys.

Table S10. Comparison of associations between any disorder and outcomes across the three surveys, by ethnicity.

Table S11. Comparison of ethnic differences over time in associations between any disorder and outcomes across the three surveys.

Table S12. Comparison of change in mental health difficulties at follow-up across the three surveys.

Table S13. Comparison of change in SDQ total difficulty score at follow-up across the three surveys, in children with any disorder at baseline.

Table S14. Comparison of associations between any emotional disorder and outcomes across the three surveys (weighted).

Table S15. Comparison of associations between any behavioral disorder and outcomes across the three surveys (weighted).

Table S16. Comparison of associations between any hyperkinetic disorder and outcomes across the three surveys (weighted).

Table S17. Comparison of associations between any comorbid disorder and outcomes across the three surveys (weighted).

Figure S1. Comparison of regression coefficients for associations between any psychiatric disorder at baseline and mental health difficulties and impact at follow-up, by age.

Acknowledgements

For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission. TND was supported by a National Institute for Health Research (NIHR) Advanced Fellowship during the preparation of this paper (NIHR300056). S.A.S.N. was supported by the Wellcome Trust (Early Career

Award 226392/Z/22/Z). All research in the Department of Psychiatry at Cambridge University benefits from the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care. The 1999 and 2004 surveys were funded through the Department of Health (now known as the Department of Health and Social Care, or DHSC), The Scottish Health Executive and the National Assembly for Wales. The 2017 survey was commissioned by NHS England on behalf of DHSC. This work was supported by the Wolfson Centre for Young People's Mental Health, established with support from the Wolfson Foundation. The authors are grateful to the Wolfson Centre Young People's Advisory Group for their input, views and advice in relation to this topic. The authors have declared that they have no competing or potential conflicts of interest.

AI statement

The Author[s] declare[s] that they have not used AI in any capacity that would require disclosure in accordance with Wiley's AI Guidelines, nor in any capacity that would reasonably require disclosure for editors, reviewers and readers to properly evaluate their research or manuscript. The author[s] confirm[s] that AI has not been used for purposes including but NOT limited to drafting and editing, to generate substantial text or restructure arguments, nor in the research methodology. The author[s] take[s] full responsibility for the accuracy of this statement.

Ethical considerations

Informed consent has been appropriately obtained. Ethical approval for secondary analysis from University of Exeter College of Medicine and Health Research Ethics Committee was obtained (Nov20/D/270). The original BCAMHS surveys had approval from Medical Research Ethics Committees (MRECs), while the original MCHYP survey had ethical approval from West London and GTAC Research Ethics Committee. The 2017 survey was reviewed and approved in April 2016 (REC reference: 16/LO/0155), with a substantial amendment approved in October 2016. The 2020 follow-up survey was approved by the National Statistician's Data Ethics Advisory Committee in May 2020.

Data availability statement

The data that support the findings of this study are available from UK Data Service. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from <https://ukdataservice.ac.uk/> with the permission of UK Data Service.

Correspondence

Tamsin Newlove-Delgado, Medical School, St Luke's Campus, University of Exeter, Heavitree Road, Exeter, EX1 2LU, UK; Email: t.newlove-delgado@exeter.ac.uk

Key points

What's known?

- The prevalence of child psychiatric disorders are increasing; however, no epidemiological study has used a clinically-validated diagnostic instrument to assess whether longer-term outcomes have improved or worsened across generations for children meeting criteria for psychiatric disorder.

What's new?

- Using standardized diagnostic assessments in three national child mental health surveys, our findings show that children with a psychiatric disorder experienced progressively worse mental health outcomes across successive generations. More than one-third of children with a disorder deteriorated in their mental health over three years in the earliest cohort, rising to over half of children in the most recent cohort.

What's relevant?

- The scale of deterioration suggests that current systems of prevention, early identification, and support are insufficient to meet children's needs. Urgent universal and public-health-oriented interventions are required to prevent worsening mental health trajectories for children with psychiatric disorders, alongside targeted clinical care.
- Our findings suggest that researchers and services must concentrate on addressing the factors that contribute to the maintenance and exacerbation of problems, as well as on the prevention, initial treatment, and provision of effective support services.

References

- Armitage, J.M., Collishaw, S., & Sellers, R. (2024). Explaining long-term trends in adolescent emotional problems: what we know from population-based studies. *Discover Social Science and Health*, 4(14), 76.
- Armitage, J.M., Kwong, A.S.F., Tseliou, F., Sellers, R., Blakey, R., Anthony, R., ... & Collishaw, S. (2023). Cross-cohort change in parent-reported emotional problem trajectories across childhood and adolescence in the UK. *The Lancet Psychiatry*, 10(7), 509–517.
- Armitage, J.M., Newlove-Delgado, T., Ford, T., McManus, S., & Collishaw, S. (2025). Characteristics of children with a psychiatric disorder in 1999, 2004 and 2017: An analysis of the national child mental health surveys of England. *Journal of Child Psychology and Psychiatry*, 66(2), 167–177.
- Armitage, J.M., Tseliou, F., Riglin, L., Dennison, C., Eyre, O., Bevan Jones, R., ... & Collishaw, S. (2023). Validation of the Strengths and Difficulties Questionnaire (SDQ) emotional subscale in assessing depression and anxiety across development. *PLoS One*, 18(7), e0288882.
- Children's Commissioner. (2021). The state of children's mental health services 2020/21. <https://assets.childrenscommissioner.gov.uk/wpuploads/2021/01/cco-the-state-of-childrens-mental-health-services-2020-21-tech-report.pdf>
- Collishaw, S., Furzer, E., Thapar, A.K., & Sellers, R. (2019). Brief report: a comparison of child mental health inequalities in three UK population cohorts. *European Child & Adolescent Psychiatry*, 28, 1547–1549.
- Collishaw, S., & Sellers, R. (2020). Trends in child and adolescent mental health prevalence, outcomes, and inequalities. *Mental Health and Illness of Children and Adolescents*, 20, 63–73.
- Copeland, W.E., Alai, I., Jonsson, U., & Shanahan, L. (2021). Associations of Childhood and Adolescent Depression With Adult Psychiatric and Functional Outcomes. *Journal of the American Academy of Child and Adolescent Psychiatry*, 60(5), 604–611.
- Ford, T., Vizard, T., Sadler, K., McManus, S., Goodman, A., Merad, S., ... & MHCYP Collaboration. (2020). Data Resource Profile: Mental Health of Children and Young People (MHCYP) Surveys. *International Journal of Epidemiology*, 49(2), 363–364.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586.
- Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well-Being Assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, 41(5), 645–655.
- Green, H., McGinnity, A., Meltzer, H., Ford, T., & Goodman, R. (2005). *Mental health of children and young people in Great Britain, 2004*. London: Palgrave Macmillan.
- Mauz, E., Houben, R., Hölling, H., & Schlack, R. (2014). Screening for Mental Health Problems in Children and Adolescents by means of the Strengths and Difficulties Questionnaire (SDQ, parental-version) within a Mixed-Mode study. *European Journal of Public Health*, 24, cku161.
- Meltzer, H., Gatward, R., Corbin, T., Goodman, R., & Ford, T. (2003). *Persistence, onset, risk factors and outcomes of childhood mental disorders*. London: TSO.
- Newlove-Delgado, T., Williams, T., Robertson, K., McManus, S., Sadler, K., Vizard, T., ... & Ford, T. (2021). *Mental Health of Children and Young People in England, 2021*. Leeds: NHS Digital.

- Parry-Langdon, N. (2008). *Three years on: Survey of the development and emotional wellbeing of children and young people*. London: Office for National Statistics.
- Riglin, L., Dennison, C., Martin, J., Tseliou, F., Armitage, J.M., Shakeshaft, A., ... & Collishaw, S. (2024). Emotional problems across development: examining measurement invariance across childhood, adolescence and early adulthood. *European Child & Adolescent Psychiatry*, 33, 4237–4245.
- Robson, E.M., Husin, H.M., Ghazaleh Dashti, S., Vijayakumar, N., Moreno-Betancur, M., Moran, P., ... & Sawyer, S.M. (2025). Tracking the course of depressive and anxiety symptoms across adolescence (the CATS study): a population-based cohort study in Australia. *Lancet Psychiatry*, 12(1), 44–53.
- Sadler, K., Vizard, T., Ford, T., Marcheselli, F., Pearce, N., Mandalia, D., et al. (2018). Mental health of children and young people in England, 2017. <https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017>
- Schlechter, P., Ford, T.J., McManus, S., & Neufeld, S.A.S. (2023). Predictors of time-varying and time-invariant components of psychological distress during COVID-19 in the U.K. Household Longitudinal Study (understanding society). *Psychological Assessment*, 35(11), 959–973.
- Sellers, R., Warne, N., Pickles, A., Maughan, B., Thapar, A., & Collishaw, S. (2019). Cross-cohort change in adolescent outcomes for children with mental health problems. *Journal of Child Psychology and Psychiatry*, 60(7), 813–821.
- StataCorp. (2021). *Stata Statistical Software: Release 17*. College Station, TX: StataCorp.
- Stuckler, D., Reeves, A., Loopstra, R., Karanikolos, M., & McKee, M. (2017). Austerity and health: the impact in the UK and Europe. *European Journal of Public Health*, 27(4), 18–21.
- Thompson, E.J., Richards, M., Ploubidis, G.B., Fonagy, P., & Patalay, P. (2023). Changes in the adult consequences of adolescent mental ill-health: findings from the 1958 and 1970 British birth cohorts. *Psychological Medicine*, 53(3), 1074–1083.
- Vizard, T., Sadler, K., Ford, T., Newlove-Delgado, T., McManus, S., Marcheselli, F., ... & Cartwright, C. (2020). *Mental health of children and young people in England, 2020*. Leeds: NHS Digital.

Accepted for publication: 1 September 2026