BMJ Open

Impact of transforming mental health services for young people in England on patient access, resource use and health: a quasi-experimental study

Stephen Rocks [1], Mina Fazel,[2] Apostolos Tsiachristas[1]

[1]Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK
[2]Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK

Correspondence to
Apostolos Tsiachristas;
apostolos.tsiachristas@ndph.ox.ac.uk

ABSTRACT

Objective To evaluate the impact of child and adolescent mental health services (CAMHS) transformation in South East England on patient access, resource utilisation and health outcomes.

Design In an observational study, we use difference-in-differences analysis with propensity score matching to analyse routinely collected patient level data.

Setting Three CAMHS services in South East England.

Participants All patients attending CAMHS between April 2012 and December 2018, with more than 57 000 spells of care included.

Main outcome measures The rate and volume of people accessing CAMHS; waiting times to the first contact and waiting times between the first and second contact; and health outcomes, including the Strengths and Difficulties Questionnaire (SDQ) and the Revised Child Anxiety and Depression Scale (RCADS).

Results The intervention led to 20% (incidence rate ratio: 1.20; 95% CI: 1.15 to 1.24) more new patients starting per month. There was mixed evidence on waiting times for the first contact. The intervention led to 10% (incidence rate ratio: 1.10; 95% CI: 1.02 to 1.18) higher waiting time for the second contact. The number of contacts per spell (OR: 1.08; 95% CI: 0.94 to 1.25) and the rereferral rate (OR: 1.06; 95% CI: 0.96 to 1.17) were not significantly different. During the post intervention period, patients in the intervention group scored on average 3.3 (95% CI: −5.0 to −1.6) points lower on the RCADS and 1.0 (95% CI: −1.8 to -0.3) points lower on the SDQ compared with the control group after adjusting for the baseline score.

Conclusions Overall, there are signs that transformation can help CAMHS achieve the objectives of greater access and improved health outcomes, but trade-offs exist among different performance metrics, particularly between access and waiting times. Commissioners and providers should be conscious of any trade-offs when undertaking service redesign and transformation.

Strengths and limitations of this study

► With little prior evidence from the evaluation of child and adolescent mental health services, this paper is a novel attempt to understand, describe and evaluate these services.
► We used data from across one large NHS Foundation Trust, which provides consistency.
► We employed state-of-the-art econometrics in a quasi-experimental design.
► However, we cannot eliminate residual confounding.
► A further limitation is data availability—specifically, health outcomes data were insufficient for including in a DiD design.

INTRODUCTION

Mental health problems impose a large health and economic burden on individuals, their families and society.[1] Globally, mental illness has been associated with reduced life expectancy and is the single biggest cause of years lived with a disability.[2 3] In the UK, it is estimated that the costs associated with mental illness totals 4.5% of gross domestic product.[4] Childhood and adolescence is crucial to any attempt to alleviate this burden. The prevalence of mental disorders in children and adolescents aged up to 18 is estimated at 13%.[5–7] Untreated mental health conditions negatively impact on development throughout the life course, making access to 'adolescent-responsive' and high-quality health systems crucially important.[8]

Child and adolescent mental health services (CAMHS) in England have come under increasing pressure and scrutiny in recent years. Between 2012/2013 and 2014/2015, there was a sharp rise in referrals to CAMHS,[9] alongside a growing number of young people presenting to emergency departments with self-harm.[10] However, over this period, many local budgets for CAMHS were reduced in real terms, resulting in reports of the perception of rising thresholds for care and limiting the number of young people able to access care, with longer waiting times in some services and considerable variation in local service provision.[11] Although effective interventions exist,[8] it is estimated that less than a quarter of young people in England with

clinically significant mental health problems receive help from specialist health services.[12]

Amid rising concerns about the adequacy of CAMHS, a 2015 UK Government review, *Future in Mind*,[13] made recommendations as to how services should change. This, along with other directives, such as the *Five Year Forward View For Mental Health*,[14] highlighted the need to 'transform' services to improve accessibility, increase quality of care and improve health outcomes. NHS England tasked Clinical Commissioning Groups (CCGs) to develop 'Local Transformation Plans' in order to access a £1.4 billion CAMHS transformation fund committed by the government to be spent from 2014/2015 to 2020/2021.

Across England, commissioners have used this additional funding to spur the local implementation of changes aligned with the transformation of CAMHS. Many CAMHS have moved from a traditional 'tiered' model, where a young person is allocated to a specific tier of service aligned with complexity/need, towards a more 'integrated' service model that allows young people to benefit from the full range of support available, recognising that needs can be complex. Other changes include new ways of working with the third sector, with some CAMHS contracting the third sector to provide support; new ways of working with schools and frequently the introduction of a single point of access (SPA) to make it easier for parents and young people to contact services.

Monitoring improvement using performance indicators is a key element of the Local Transformation Plans and a prerequisite for accessing funding. Although these indicators provide commissioners with information about trends in time, they cannot provide conclusive evidence about whether CAMHS transformation has been successful in achieving its objectives. With only one financial year left to spend the transformation funds, evidence is needed to inform future investment in CAMHS and any adaptation or redesign of transformation strategies.

In this study, we investigated the impact of the transformation of CAMHS in Buckinghamshire compared with neighbouring CAMHS in Oxfordshire and Swindon, Wiltshire, Bath and North East Somerset (collectively referred to as SWBaNES) on patient access, resource use and health outcomes, using a series of service and individual level outcomes. We used data for all patients and contacts with CAMHS between 2012/2013 and 2018/2019. The study is part of a larger mixed methods evaluation.[15]

METHODS
Study design
The development and implementation of local transformation plans occurred according to a range of factors, such as local decision-making and capacity. Since the timing and content of the transformation was not subject to randomisation, we adopted a quasi-experimental study design by using difference-in-differences (DiD) analysis with propensity score matching (PSM) to reduce observed and unobserved bias.[16 17] This type of study

design is suggested by the Medical Research Council in England, among others, to overcome the significant challenges inherent in the evaluation of complex health interventions.[18–21]

Setting
The study was conducted in the South of England (online supplementary figure 1) and included CAMHS provided by Oxford Health NHS Foundation Trust (Oxford Health). Oxford Health is one of the largest CAMHS providers in England. Delivering services for NHS Buckinghamshire CCG, NHS Oxfordshire CCG, NHS Swindon CCG, NHS Wiltshire CCG and NHS Bath and North East Somerset CCG (NHS BaNES CCG). As the latter three are commissioned under one contract, they are referred to collectively as SWBaNES. For more information on the study setting, see online supplementary table 1.

Intervention and control groups
The transformation of CAMHS across the study geography differed in intensity, timing and pace. Transformation started in 2015/2016 in Buckinghamshire (intervention). The core components of transformation were similar across the sites. All services adopted some version of the THRIVE model.[22] This involved a shift towards treatment decisions based on level of need, or capacity to benefit, rather than specific diagnoses or interventions. It required redesigning pathways, including introducing 'Getting Help' and 'Getting More Help' pathways broadly in place of tiers 2 and 3, and a redesignation of staff. Another core shift was the introduction of an SPA for referrals. In conventional services, primary care, in the form of general practitioners (GPs), acts as the gatekeeper to CAMHS. The introduction of an SPA, including the ability for young people (primarily aged 16–17) to self-refer, was intended to make it easier for young people and parents to engage with CAMHS by allowing them to contact services directly. A final core element of transformation was new ways of working with the third sector. In Buckinghamshire, third sector involvement was primarily with Barnardo's, one of the largest charities providing care to children in England. Barnardo's workers were placed into the SPA as Contact Support Workers where they would take initial calls and liaise with clinicians about referral decisions; and into the new 'Getting Help' pathway as 'Buddies'. Buddies were trained to provide a low-intensity intervention of up to six sessions for children and young people with low-to-moderate difficulties.

While recognising that service transformation is a process rather than a discrete event, we dated the transformation to the financial year in which it started in the intervention group, 2015/2016. The study therefore has a pretransformation and post-transformation period.

Transformation did not start in Oxfordshire and SWBaNES until 2018/2019 (control). In the interim, the control group continued to deliver services based on a 'traditional, tiered system' in which CAMHS were

accessed primarily through GPs with some referrals from school staff and social services. Many of the goals and the core components of transformation were similar in Oxfordshire and SWBaNES to those in Buckinghamshire, although some aspects of delivery were shaped according to local assets and capacity (eg, the role of the third sector).

Data

Data on referrals to CAMHS, contacts (ie, appointments) with CAMHS, diagnosis and health outcomes of all patients who started their contact with CAMHS between April 2012 and December 2018 were obtained from the electronic patient record systems of Oxford Health. Data prior to April 2012 were incomplete and were not considered valuable for the purposes of this study. The Strengths and Difficulties Questionnaire (SDQ) and the Revised Children's Anxiety and Depression Scale (RCADS) were the only widely recorded health outcomes available. Health outcome data were available from 2014, but recording was very limited until 2015. We also obtained sociodemographic data including gender, ethnicity, age at referral to CAMHS and socioeconomic deprivation based on the postcode of patient's residence or patient's GP practice (which is most often based on the residence). The data were grouped in distinct spells of care. A spell of care was defined as an overlapping bundle of referrals to and contacts with CAMHS. Specifically, based on clinical input, we considered a referral to be overlapping if the new referral date fell within 28 days weeks of the previous discharge date. More detail is provided in online supplementary figure S2.

Outcomes

A set of outcomes was defined to assess the impact of the CAMHS transformation on patient access, resource use and health (table 1). For access to care, the rate of new patients per month per 1000 people aged under 18 was used to capture overall access to the service. This was the only outcome defined at service level (ie, Buckinghamshire—intervention; Oxfordshire, SWBaNES—control).

In addition, we analysed waiting times from referral to the first contact with CAMHS and waiting times from the first to the second contact. In total, 5% of spells did not have any attended contacts and 19% of spells did not have a second contact. We further excluded 189 spells with a waiting time to the first contact of more than 355 days (<1%) and 437 spells (<1%) with a waiting time for a second contact over 235 days. This was based on visual inspection of the distribution of the data which revealed a long tail of extreme values and was further tested using Cook's Distance, according to which all values over the respective number of days were considered extreme.

With regard to resource use, two outcomes were used: the number of contacts per spell and the rate of rereferrals—a binary outcome recording whether a patient has more than once distinct spell of care (the first spell is coded 0, any further spells 1). Furthermore, the total scores on the SDQ (range: 0–40) and RCADS (range: 0–141) questionnaires were used as health outcomes. Both RCADS and the SDQ can be completed by the child or parent. We created a single measure for each that combined the child and parent scores, giving priority to the child response where both were recorded and controlling for respondent type as an additional covariate.

Propensity score matching

We used PSM to account for observed differences between the intervention and control group that would influence the treatment effect and for changes in the composition of the groups over time. Specifically, when analysing spell level outcomes, we accounted for potential differences by estimating a multigroup propensity score to weight each spell.[23] Multinomial logistic regression was used to predict the probability of each spell being in one of four groups (precontrol, postcontrol, preintervention and postintervention) as a function of observed covariates. We then ascribed each spell a weight according to its likelihood of being in the intervention group in the preperiod. For spell i, the weight was calculated using the following formula:

Table 1	Outcomes		
Objective	**Outcome**	**Measure**	**Level**
Patient access	Rate of new patients per month	Restricted to new patients only, a count of spells starting per month expressed per 1000 under 18	Service
	Waiting times: first contact	Days between referral received and first attended contact	Spell
	Waiting times: second contact	Days between first attended contact and next attended contact	Spell
Resource use	Number of attended contacts per spell	Number of attended contacts within a spell	Spell
	Rereferral rate	Binary outcome indicating whether a spell is followed by a subsequent spell (rereferral)	Spell
Health	Revised Child Anxiety and Depression Scale (RCADS)	Last outcome measure when there are least two measures recorded (RCADS)	Spell
	Strengths and Difficulties Questionnaire (SDQ)	Last outcome measure when there are least two measures recorded (SDQ)	Spell

RCADS, Revised Child Anxiety and Depression Scale; SDQ, Strengths and Difficulties Questionnaire.

$$w_i = e_1\left(x_i\right) / e_g\left(x_i\right)$$

where the intervention group in the preperiod is denoted 1 and the group the spell starts in is denoted g. Weights were included in the regression models outlined above to balance the selected covariates across the four groups. Given the limited recording of health outcomes, we performed the PSM separately for the SDQ and RCADS. Since data on health outcomes were almost entirely restricted to the postperiod, there were only two groups (control and intervention).

In selecting the covariates for inclusion, we sought to avoid post-treatment bias by excluding factors that we suspected the intervention would directly address.[24] We therefore included only demographic variables, rather than service factors such as the level of care or urgency of referral. The following covariates featured in our analysis of all outcomes: gender (male/female), ethnicity (White British/White Other/Black, Asian and minority ethnic), age at referral received and deprivation index. When analysing health outcomes, we included the baseline score from the respective health questionnaire to control for differences in initial severity. Among our covariates, ethnicity was the only major source of missing data (n=9963, 17.3%). To account for the missing data in the analysis, we used multiple imputation. Our approach is detailed in the online supplementary table S2.

For the PSM, covariate balance was assessed based on the mean and median percentage standardised bias as well as Rubin's B (the absolute standardised difference of the means of the linear index of the propensity score in the treated and (matched) non-treated group) and Rubin's R (the ratio of treated to (matched) non-treated variances of the propensity score index).[25] Following Rubin's recommendation,[26] we considered B less than 25 and R between 0.5 and 2 to indicate sufficient balance.

Regression analysis
We included all of the same covariates in our regression analysis as for PSM: gender, ethnicity, age at receipt of referral, deprivation and, for health outcomes only, initial severity. When analysing data at a service level, we took the average for deprivation and age and calculated proportions for other covariates, specifically the proportion male and the proportion White British. At the service level we also included the quarter the spell started in to account for any seasonality.

Following common practice, we first specified a DiD model that assumes parallel trends in the outcomes between the intervention and control group (model 1).

$$Y_{it} = \beta_0 + \beta_1 * Int_i + \beta_2 * TP_t + \beta_3 * Int_i TP_t + \beta_4 * X_{it} + \beta_5 * Q_j + \varepsilon_{it} \qquad (1)$$

In equation 1, the dependent variable Y_{it} denotes one of the nine outcomes for spell (or service) i at time t. Int_i denotes the intervention group (taking the value 1 for the intervention group and 0 for the control group) and TP_t accounts for the transformation period (taking the value

0 before transformation and 1 after transformation). The interaction term $Int_i TP_t$ is the variable of interest (ie, the difference-in-difference estimator) and captures the impact of the transformation on the outcome. The vector X_{it} includes the covariates described above, and Q_j, a time variable indicating the quarter of each year and capturing any linear time trends. The error term ε_{it} captures unobserved variations in the outcome.

To relax the parallel trends assumption, which is critical for producing unbiased DiD estimates,[23] we specified a model with differential trends between the intervention and control group (model 2). This model included the term, Q_{aj}, which accounts for different linear trends across the service areas in the study.

$$Y_{it} = \beta_0 + \beta_1 * Int_i + \beta_2 * TT_t + \beta_3 * Int_i TT_t + \beta_4 * X_{it} + \beta_5 * Q_j + \beta_6 * Q_{aj} + \varepsilon_{it} \qquad (2)$$

To test whether the parallel assumption holds, we limited the analysis period to observations in the pretransformation period (ie, before 2015/2016), assigned an arbitrary date in 2013/2014 as the pseudotransformation date, and ran model 1. If the DiD estimator was statistically significant in this test, the parallel trends assumption was violated.

Finally, we were unable to employ a DiD design for health outcomes. Recording of health outcomes was only recently introduced in CAMHS and we had insufficient data to form a reliable preperiod. Instead, we performed a regression analysis with PSM to compare the health outcomes measures at follow-up (ie, the latest recorded health outcome of a patient in a spell) between the intervention and control group adjusting for differences at baseline measures (ie, the first recorded health outcome in the same spell), using the following model:

$$Y_{it} = \beta_0 + \beta_1 * Int_i + \beta_2 * X_{i0} + \beta_3 * Q_j + \varepsilon_{it} \qquad (3)$$

where Y_{it} is the last outcomes score. The vector, X_{i0}, includes a range of covariates defined at baseline including the first score on the outcome scale. The variable of interest here is $\beta_1 * Int_i$.

All regressions were fitted using generalised linear models with clustered standard errors at patient level (or service level for the service level outcome) and included the weights from the PSM. The distribution family in each regression models was chosen based on a Modified Park Test (using the user-command 'GLMdiag' in Stata) and the Akaike and Bayesian information criteria. The link function was selected with reference to the link test ('linktest' command in Stata).

Sensitivity analysis
In order to assess the sensitivity of the results to the specified pretreatment period, we replicated the analysis with a different timeframe. While in the main analysis we utilised all complete data for our preperiod and postperiod, in our sensitivity analysis, we followed the approach of Slaughter in assessing the robustness of the results to a shortened pretreatment period.[27] In our case,

we restricted the data to starting from October 2013. We specified this time point following visual inspection of trends in the preperiod (see online supplementary figures S3 and S4).

RESULTS

Observed differences between intervention and control

There were 57 501 spells of care initiated between April 2012 and December 2018. They differed between the intervention and control group in terms of sociodemographic characteristics (table 2). In the preintervention period, the intervention group included younger patients and patients from less socioeconomically deprived areas, and consisted of proportionally more males and fewer patients with White British ethnicity. These differences were similar in the postintervention period.

Table 2 also shows the change in the outcomes from the preperiod to postperiod. The rate of new patients starting per month per 1000 under 18 increased in the intervention group (from 1.05 to 1.22), but not in the control (1.24 in both). Waiting times for the first contact were lower in the intervention group (39 days) than the control (41 days) in the preperiod. This gap widened in the postperiod, with the wait remaining 39 days for the intervention, but increasing to 61 days for the control. By contrast, waits for the second contact were higher in the intervention group (27 days vs 23 days for the control). This gap also increased in the postperiod (35 days vs 27 days). Attended contacts per spell were higher in the intervention group (17.0 contacts) in the preperiod compared with the control (15.8). In the postperiod, the gap narrowed (10.2 in the intervention compared with 10.4 in the control group). The difference in the rereferral rate was consistent in the preperiod and postperiods. In the preperiod, the re-referral rate was higher in the intervention group (0.35) than the control (0.32); the difference narrowed in the postperiod, at 0.17 for the intervention and 0.16 in the control. In terms of health outcomes, the last RCADS score was 4.6 points lower in the intervention group and the SDQ score was 1.2 points lower.

Achieved covariance balance

Following PSM, observed confounding was reduced to acceptable levels according to Rubin's B (ie, it was lower than 25 in the preperiod and postperiods) and Rubin's R (ie, it was between 0.5 and 2 in the preperiod and postperiod) across all imputed data sets. Likewise, the mean and median percentage standardised biassed were also substantially reduced to values closer to 0. The results of the PSM in terms of covariate balance from one imputed data set are presented in table 3 and illustrated graphically in the online supplementary figure 5. The results across each imputed data set were very similar.

Results from the regression analysis

Table 4 provides the results of the difference-in-difference analysis with PSM under the selected model, as well as whether the parallel trends assumption held.

The intervention led to 20% (incidence rate ratio: 1.20; 95% CI: 1.15 to 1.24) more new patients starting per month. Waiting times were lower in the intervention group under the parallel trends model, however the parallel trends assumption was violated. With differential trends, the waiting times for the first contact instead increased 83% (incidence rate ratio: 1.83; 95% CI: 1.67 to 1.99). The intervention led to 10% (incidence rate ratio: 1.10; 95% CI: 1.02 to 1.18) higher waiting time for the second contact with CAMHS compared with the control group. The parallel trends assumption was also violated for the number of contacts per spell. Under the differential trends model, the point estimate was 1.08, but inconclusive (OR: 1.08; 95% CI: 0.94 to 1.25). The rereferral rate was unchanged following the intervention (OR: 1.06; 95% CI: 0.96 to 1.17). During the post intervention period, patients in the intervention group scored on average 3.3 (95% CI: −5.0 to −1.6) points lower on the RCADS and 1.0 (95% CI: -1.8 to -0.3) points lower on the SDQ compared with the control group after adjusting for the baseline score. Table 4 also shows the distribution family and link function that provided the best goodness-of-fit to the data. The results for each model and from the test of the parallel trends assumption are presented in the online supplementary table 3, and graphical visualisation of the observed trends in outcomes is included in online supplementary figures S3 and S4.

Sensitivity analysis

Table 5 shows the results of the sensitivity analysis. The intervention led to 16% (incidence rate ratio: 1.16; 95% CI: 1.13 to 1.20) more spells starting per month. Waiting times for the first contact fell 14% (incidence rate ratio: 0.86; 95% CI: 0.81 to 0.91), but rose for the second contact compared with the control group (incidence rate ratio: 1.11; 95% CI: 1.03, to 1.19). The parallel trends was again violated for the number of attended contacts per spell; under the differential trends model, the number of attended contacts was 12% higher (OR: 1.12; 95% CI: 0.96, 1.31). The re-referral rate was unchanged (OR: 1.03; 95% CI: 0.91 to 1.15). The results for each model and from the test of the parallel trends assumption are presented in the online supplementary table S4.

DISCUSSION

This empirical evaluation provides the first evidence about the impact of CAMHS transformations in South England on access to care, service use and health outcomes. There are mixed signs on access to care, resource use was unchanged and there is tentative evidence that health outcomes are better under a transformed service.

There is increased access to care in the transformed CAMHS relative to the control as measured by the rate of new patients accessing care. However, waiting times for the second contact increased following transformation. For waiting times for the first contact, our various specifications produced widely different estimates for the effect,

Table 2 Outcomes and covariates for intervention and control, preintervention and postintervention (unweighted)

| | Preintervention | | | | | Postintervention | | | | | |
| | Intervention | | Control | | (1–2) | Intervention | | Control | | (1–2) | |
	N/(clusters)	% or Mean/(SD)	N/(clusters)	% or Mean/(SD)	Diff.	N/(clusters)	% or Mean/(SD)	N/(clusters)	% or Mean/(SD)	Diff.
Covariates										
Percent male (%)	5116 (4497)	49.75	17338 (15053)	44.85	4.90***	9341 (7811)	48.24	25748 (21938)	44.91	3.33*
Percent White British (%)	4359 (3784)	84.06	15674 (13536)	91.31	−7.25***	6518 (5257)	81.74	21014 (17641)	88.54	−6.79***
Mean deprivation (Ward index of multiple deprivation)	5116 (4497)	10.78 (7.32)	17338 (15053)	13.64 (8.82)	−2.86***	9341 (7811)	10.90 (7.99)	25748 (21938)	13.47 (8.94)	−2.57***
Mean age at referral received	5116 (4497)	12.70 (3.97)	17338 (15053)	13.13 (5.49)	−0.43***	9341 (7811)	12.49 (4.30)	25748 (21938)	13.16 (5.43)	−0.66***
Outcomes										
Rate of new patients per month	36 (1)	1.05 (0.00)	72 (2)	1.24 (1.92)	−0.20	45 (1)	1.22 (0.00)	90 (2)	1.24 (2.64)	−0.02
Waiting time: first contact	4803 (4253)	39.49 (43.81)	16608 (14528)	41.20 (48.78)	−1.71**	8765 (7450)	38.95 (51.25)	24274 (20859)	60.86 (66.63)	−21.91***
Waiting time: second contact	3985 (3611)	27.46 (36.06)	14029 (12563)	23.08 (32.32)	4.38***	6677 (5871)	34.88 (44.18)	18969 (16836)	26.54 (35.78)	8.33***
Attended contacts per spell	5116 (4497)	16.97 (33.44)	17337 (15052)	15.78 (31.51)	1.19**	8630 (7204)	10.23 (15.67)	25283 (21508)	10.41 (19.20)	−0.18
Rereferral rate	5116 (4497)	0.35 (0.48)	17338 (15053)	0.32 (0.48)	0.03***	9341 (7811)	0.17(0.37)	25748 (21938)	0.16 (0.36)	0.02***
Health outcomes: RCADS	Not available					2323 (1552)	53.89 (38.01)	4154 (2703)	58.53 (40.12)	−4.64***
Health outcomes: SDQ	Not available					1227 (799)	40.68 (13.39)	3171 (1951)	41.85 (13.26)	−1.16***

The column (1–2) displays the differences across the groups. ***, ** and * indicate significance at the 1%, 5% and 10% critical level, tested using a t-test for continuous data and a χ^2 test for categorical variables. SEs are clustered at the patient or service level (service level includes all three sites in the study: Buckinghamshire, Oxford, SWBaNES).

RCADS, Revised Child Anxiety and Depression Scale ; SDQ, Strengths and Difficulties Questionnaire.

Table 3 Covariate balance: matched versus unmatched, preintervention and postintervention

		Mean bias	Median bias	B	R
Pre	Unmatched	20.2	15.6	49.1*	0.85
	Matched	1.2	0.9	3.7	1.28
Post	Unmatched	18.8	16.4	44.4*	0.81
	Matched	2.3	2.6	6.1	0.86
RCADS only					
Post	Unmatched	15.2	13.6	40.6*	1.02
	Matched	1.6	0.6	4.8	1.34
SDQ only					
Post	Unmatched	11.6	5.2	36.7	0.96
	Matched	1.0	0.7	2.9	1.36

*if B>25%, R outside [0.5; 2].
RCADS, Revised Child Anxiety and Depression Scale ; SDQ, Strengths and Difficulties Questionnaire.

with a decrease under the parallel trends model but a large increase when adjusting for differential trends. Visual inspection confirms divergent trends in the earliest periods of the study: prior to October 2013 waiting times for the first contact fell in the intervention group while rising for the control group, after this point both groups exhibit a similar trend (see online supplementary figure S3). In our sensitivity analysis, from October 2013 onwards, the trends were parallel and we found a fall in waiting times for the first contact. Given the inflexion in waiting times observed in the intervention group in the post-treatment period after a steadily rising trend in the immediate pretreatment periods, it seems plausible that the intervention reduced waiting times in the intervention group. Yet, it is also possible that an earlier reduction in waiting times—prior to October 2013—in the

intervention group is one reason for the earlier adoption of the transformation plan in Buckinghamshire, which raises doubts as to whether other areas would be able to replicate this success. Taken together, this amounts to evidence of an increase in waiting times for the second contact alongside more people accessing the service, but greater uncertainty over the impact on waiting times for the first contact.

In terms of resource use, the rereferral rate was unchanged. There was also no significant change in the overall number of attended contacts per spell. Again, the results diverge depending on the model, with a reduction in contacts per spell found in the parallel trends model, although this was also inconclusive. A fall in attended contacts per patient would be expected under the Thrive model, which looks to curtail over-lengthy treatment. Our results again hint that Buckinghamshire, as an early adopter of the transformation, may have differed from the control group. However, contacts per spell as an outcome could be explored in greater detail. Data from this study reveal that some young people attend CAMHS only once and many patients are seen less than three times as part of a spell of care.[28] The average says little about changes in the pattern of use among frequent or infrequent attenders and future research could seek to understand whether changes exist at different levels of service use. This paper is part of a wider mixed-methods evaluation that may be able to shed further light on this.[15]

Finally, this study also finds evidence that patients leave with statistically lower scores on the two main health outcomes questionnaires—SDQ and RCADS—when controlling for the baseline score and the type of questionnaire. This suggests higher quality of life in the intervention group. However, the evidence is weaker since we have much less comprehensive recording and insufficient data in the preperiod to test this in a DiD framework. It could be that the intervention and control groups have

Table 4 Main results

	Rate of new patients per month	Waiting time for first contact	Waiting time for second contact	Attended contacts per spell	Rereferral rate	RCADS: last measure	SDQ: last measure
Model	PT (1)	DT (2)	PT (1)	DT (2)	PT (1)	(3)	(3)
DiD	1.20***	1.83***	1.10*	1.08	1.06	−3.3	−1.0**
SE	(0.02)	(0.08)	(0.04)	(0.08)	(0.05)	(0.87)	(0.39)
95% CI	(1.15 to 1.24)	(1.67 to 1.99)	(1.02 to 1.18)	(0.94 to 1.25)	(0.96 to 1.17)	(−5.0 to −1.6)	(−1.8 to −0.3)
Parallel trends	Yes	No	Yes	No	Yes	n/a	n/a
Family	Poisson	Poisson	Poisson	Gamma	Binomial	Gaussian	Gaussian
Link	Log	Log	Log	Log	Logit	Identity	Identity
Observations	243	54 450	44 097	56 366	57 543	6484	4401

The coefficients of the regressions with log link function are reported in the exponential form; the outcome rate of new patients per month was specified at service level.
*p-value<0.001, **p-value<0.05, and *p-value<0.01.
DT, differential trends; n/a, not applicable; PT, parallel trends; RCADS, Revised Child Anxiety and Depression Scale ; SDQ, Strengths and Difficulties Questionnaire.

Table 5 Sensitivity analysis—restricted preperiod (starting October 2013)

	Rate of new patients per month	Waiting time for first contact	Waiting time for second contact	Attended contacts per spell	Rereferral rate
Model	PT (1)	PT (1)	PT (1)	DT (2)	PT (1)
DiD	1.16***	0.86***	1.11**	1.12	1.03
SE	(0.02)	(0.02)	(0.04)	(0.09)	(0.06)
95% CI	(1.13 to 1.20)	(0.81 to 0.91)	(1.03 to 1.19)	(0.96 to 1.31)	(0.91 to 1.15)
Parallel trends	Yes	Yes	Yes	No	Yes
Family	Poisson	Poisson	Poisson	Gamma	Binomial
Link	Identity	Log	Log	Log	Logit
Observations	189	44 931	35 486	46 431	47 608

The coefficients of the regressions with log link function are reported in the exponential form; the outcome rate of new patients per month was specified at service level.
*p-value<0.001, **p-value<0.05, and *p-value<0.01.
DT, differential trends; PT, parallel trends.

been consistently different prior to the intervention, but we cannot test for this. The differences reported are also at the average; future research could look more closely at final outcomes scores in order to understand the clinical significance of this shift. Improved collection of routine outcomes data will remain important for the evaluation of CAMHS in the future.

Overall, this analysis highlights some of the trade-offs inherent in changing a complicated organisation such as CAMHS. We found an increase in activity in the intervention group relative to other services, which would be expected given the increased funding associated with the transformation. Such overall increases in access are needed to meet recently introduced targets on the numbers of people estimated to have a mental health condition receiving care.[14] However, attempts to improve one may have consequences elsewhere in the system. Such trade-offs have been illustrated with reference to, for instance, liaison mental health services.[29] In particular, this study finds evidence of an increase in access alongside an increase in the waiting time for a second contact, suggesting the increased workload brought about by the transformation was not fully matched by sufficient capacity further along the system. Commissioners and providers should work together to anticipate such changes but must also consider the relative importance of different outcomes.

Strengths
Strengths of this study include the use of all data across one large NHS Foundation Trust. The scale of the research means the analyses are highly powered. With little prior evidence from the evaluation of CAMHS, this paper is a novel attempt to understand, describe and evaluate these services. An advantage of having data from one trust is that the recording system is consistent, increasing the comparability of the data. We employed a DiD framework, as recommended by MRC among others.[16 19] Examples of the use of a DiD framework to evaluate service level interventions include stroke unit centralisation and

policy changes such as the introduction of payment by results or liberalising bar closure hours.[30–35] An additional concern with repeated cross-sectional analysis is that the composition of the populations may evolve over time, exogenous to the intervention, biassing results. Given this risk, we combined the DiD framework with PSM to reduce the bias in our estimates.[23 36 37] A number of studies have applied this approach, including evaluations of NHS Health Checks, multidisciplinary teams and extended access to primary care.[38–40]

Limitations
Limitations of the study include residual confounding. We have applied techniques appropriate to the non-experimental design, but cannot entirely eliminate the potential for residual confounding.[41] A further issue is data availability, most particularly the lack of data on health outcomes or patient experience—specifically, health outcomes data were insufficient for including in a DiD design. We were also limited in the number of available comparator groups, with a greater number of CAMHS as controls we could employ other techniques, such as synthetic matching which would permit us to match on outcomes in the preperiod.[42] Although, trends appeared broadly consistent across the groups, the major exception was waiting times for the first contact. We sought to control for this using a differential trend, although this may be too restrictive an assumption.[43] We also conducted a sensitivity analysis restricting the pretreatment period. Future research could explore a more flexible difference-in-differences model to better understand the sensitivity of the estimate to pretreatment dynamics.[44] Finally, the intervention took place in a relatively affluent part of England, raising questions as to the extent to which results can be extrapolated for other CAMHS in the country.

Implications
This study looked at one of the early adopters of CAMHS transformations. As such, the findings should be used to

help steer transformations going forward. In addition, there is the need for embedded monitoring of transformations informed by richer data and detail of the transformation objectives. The set of techniques proposed in this study could be applied to evaluate the impacts of transformation elsewhere where the same opportunity exists relating to differential adoption of transformation plans, and also to other service level interventions at various levels.

CONCLUSIONS

This study provides the first robust evidence of the impact of CAMHS transformation in England. There are signs that transformation can help CAMHS achieve the objectives of greater access and improved health outcomes, but trade-offs exist among different performance metrics, particularly between access and waiting times. Commissioners and providers should be conscious of any trade-offs when undertaking service redesign and transformation. Improved capture of routine patient outcomes and experience would enhance our understanding of the differential impact of the various components of the transformation.

Patient and Public Involvement

We conducted extensive Patient and Public Involvement (PPI) activities ahead of this study. We presented our plans and questions to the Oxon Young Person's advisory group (25 members) who gave feedback on the current transformation plan, in particular around how they wanted to access services, concerns about privacy when being seen outside of traditional CAMHS settings and about seeing non-CAMHS professionals. These same questions were asked to a further 15 secondary students. In addition, 60 teachers at different schools were asked what they and their students with mental health problems needed from services. This helped inform our understanding of the context for changes in services.

ORCID iD
Stephen Rocks http://orcid.org/0000-0001-7633-7708

REFERENCES

1 Doran CM, Kinchin I. A review of the economic impact of mental illness. *Australian health review: a publication of the Australian Hospital Association* 2017. [Epub ahead of print: 14 Nov 2017].
2 Chesney E, Goodwin GM, Fazel S. Risks of all-cause and suicide mortality in mental disorders: a meta-review. *World Psychiatry* 2014;13:153–60.
3 Vigo D, Thornicroft G, Atun R. Estimating the true global burden of mental illness. *The Lancet Psychiatry* 2016;3:171–8.
4 OECD. *Mental health and work: United Kingdom*, 2014.
5 Kessler RC, Berglund P, Demler O, *et al*. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National comorbidity survey replication. *Arch Gen Psychiatry* 2005;62:593–602.
6 NHS Digital. Mental health of children and young people in England, 2017, 2018. Available: https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2017/2017 [Accessed 01 Jul 2019].
7 Polanczyk GV, Salum GA, Sugaya LS, *et al*. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J Child Psychol Psychiatry* 2015;56:345–65.
8 Neufeld SAS, Dunn VJ, Jones PB, *et al*. Reduction in adolescent depression after contact with mental health services: a longitudinal cohort study in the UK. *The Lancet Psychiatry* 2017;4:120–7.
9 Earle DJ. Children and young people's mental health British Medical Association 2016.
10 Hawton K, Saunders KEA, O'Connor RC. Self-Harm and suicide in adolescents. *The Lancet* 2012;379:2373–82.
11 Frith E. Progress and challenges in the transformation of children and young people's mental health care: a report of the education policy Institute's mental health Commission: education policy Institute 2016.
12 Kessler RC, Amminger GP, Aguilar-Gaxiola S, *et al*. Age of onset of mental disorders: a review of recent literature. *Curr Opin Psychiatry* 2007;20:359–64.
13 Department of Health. Future in mind: promoting, protecting and improving our children and young people's mental health and wellbeing. gov.uk 2015.
14 NHS England. The five year forward view for mental health 2016.
15 Rocks S, Stepney M, Glogowska M, *et al*. Understanding and evaluating new models of child and adolescent mental health services in south-east England: a study protocol for an observational mixed-methods study. *BMJ Open* 2018;8:e024230.
16 Raine R, Fitzpatrick R, Barratt H, *et al*. Health services and delivery research. challenges, solutions and future directions in the evaluation of service innovations in health care and public health. *Southampton* 2016.
17 Bell SH, Olsen RB, Orr LL, *et al*. Estimates of external validity bias when impact evaluations select sites nonrandomly. *Educ Eval Policy Anal* 2016;38:318–35.
18 Sutton M G-BS, Martin G, *et al*. Economic analysis of service and delivery interventions in health care: National Institute for health research 2018.
19 Craig P, Dieppe P, Macintyre S, *et al*. Developing and evaluating complex interventions: the new medical Research Council guidance. *BMJ* 2008;337:a1655.
20 Anderson R. New MRC guidance on evaluating complex interventions. *BMJ* 2008;337:a1937.
21 Craig P, Petticrew M. Developing and evaluating complex interventions: reflections on the 2008 MRC guidance. *Int J Nurs Stud* 2013;50:585–7.
22 Wolpert MHR, Hodges S, *et al*. *Thrive elaborated*. 2nd edition. London: CAMHS Press, 2016.

23 Stuart EA, Huskamp HA, Duckworth K, *et al*. Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Serv Outcomes Res Method* 2014;14:166–82.

24 Rosenbaum PR. The Consquences of adjustment for a concomitant variable that has been affected by the treatment. *J R Stat Soc Ser A* 1984;147:656–66.

25 Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* 1985;39:33–8.

26 Rubin DB. Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Servic Outcomes Res Methodol* 2001;2:169–88.

27 Slaughter MJ. Trade liberalization and per capita income convergence: a difference-in-differences analysis. *J Int Econ* 2001;55:203–28.

28 Wolpert M, Harris R, Jones M, *et al*. *Thrive: the AFC-Tavistock model for CAMHS*, 2014.

29 House A, Guthrie E, Walker A, *et al*. A programme theory for liaison mental health services in England. *BMC Health Serv Res* 2018;18:742.

30 Morris S, Hunter RM, Ramsay AIG, *et al*. *Impact of centralising acute stroke services in English metropolitan areas on mortality and length of hospital stay: difference-in-differences analysis*. 349. BMJ: British Medical Journal, 2014.

31 Morris S, Ramsay AIG, Boaden RJ, *et al*. Impact and sustainability of centralising acute stroke services in English metropolitan areas: retrospective analysis of hospital episode statistics and stroke national audit data. *BMJ* 2019;11:l1.

32 Friebel R, Hauck K, Aylin P. Centralisation of acute stroke services in London: impact evaluation using two treatment groups. *Health Econ* 2018;27:722–32.

33 Green CP, Heywood JS, Navarro M. Did liberalising bar hours decrease traffic accidents? *J Health Econ* 2014;35:189–98.

34 Round J, Drake R, Kendall E, *et al*. Evaluating a complex system-wide intervention using the difference in differences method: the delivering choice programme. *BMJ Support Palliat Care* 2015;5:26–33.

35 Farrar S, Yi D, Sutton M, *et al*. Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis. *BMJ* 2009;339:b3047.

36 Wing C, Simon K, Bello-Gomez RA. Designing difference in difference studies: best practices for public health policy research. *Annu Rev Public Health* 2018;39:453–69.

37 Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med* 2015;34:3661–79.

38 Whittaker W, Anselmi L, Kristensen SR, *et al*. Associations between extending access to primary care and emergency department visits: a Difference-In-Differences analysis. *PLoS Med* 2016;13:e1002113.

39 Chang KC-M, Lee JT, Vamos EP, *et al*. Impact of the National health service health check on cardiovascular disease risk: a difference-in-differences matching analysis. *Can Med Assoc J* 2016;188:E228–38.

40 Stokes J, Kristensen SR, Checkland K, *et al*. Effectiveness of multidisciplinary team case management: difference-in-differences analysis. *BMJ Open* 2016;6:e010468.

41 O'Neill S, Kreif N, Grieve R, *et al*. Estimating causal effects: considering three alternatives to difference-in-differences estimation. *Health Serv Outcomes Res Method* 2016;16:1–21.

42 Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *J Am Stat Assoc* 2010;105:493–505.

43 Mora R, Reggio I. Alternative diff-in-diffs estimators with several pretreatment periods. *Econom Rev* 2019;38:465–86.

44 Mora Villarrubia R, Reggio I. didq: a command for treatment-effect estimation under alternative assumptions. *Stata J* 2015;15:281132.