medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Title

Effect of a needs-based model of care on the characteristics of healthcare services in England: the i-THRIVE National Implementation Programme

Short title

Effect of needs-based care on healthcare services

Authors

R Sippy[1], L Efstathopoulou[2,3], E Simes[2,3], M Davis[2,3], S Howell[2,3], B Morris[2], O Owrid[2,3], N Stoll[2,3], A Moore[1,2,3]*, P Fonagy[2,3]

[1]Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom

[2]Anna Freud National Centre for Children and Families, London, United Kingdom

[3]Department of Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom

* corresponding author, am2708@medschl.cam.ac.uk, Clifford Allbutt Building, Hills Road, Cambridge, CB2 0XY

Word count

3,584

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Aims: Developing integrated mental health services that focus on the needs of children and young people is a key policy goal in England. The THRIVE Framework and its associated implementation programme, i-THRIVE, are now used in areas covering over 65% of England's children. This study explores the experiences of staff involved with the i-THRIVE programme, assesses its effectiveness, and examines how local system working relationships influence the programme's success.

Methods*:* The i-THRIVE programme was evaluated among twenty participating sites (ten implementation and ten comparison sites). Measurements included surveys of staff and leaders at each site and assessment of the “THRIVE-like” features of each site. Additional site-level characteristics were collected from health system reports. The effect of i-THRIVE was evaluated using a four-group propensity-score weighted difference-in-differences model; the moderating effect of local system working relationships was evaluated with a difference-in-difference-in-differences model.

Results: The results show that strong working relationships in the local system significantly enhances the effectiveness of the i-THRIVE programme. Sites with highly effective working relationships showed a notable improvement in "THRIVE-like" features, with an average increase of 16.41 points (95% confidence interval: 1.69–31.13, p-value: 0.031) compared to control sites. In contrast, sites with ineffective working relationships did not benefit from the i-THRIVE programme (-2.76, 95% confidence interval: -18.25–12.73, p-value: 0.708). This influence of working relationship effectiveness was consistent across various levels of THRIVE features.

Conclusions*:* The findings underscore the importance of working relationship effectiveness in the successful adoption and implementation of health policies like i-THRIVE.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Background

Mental healthcare has numerous well-supported models. However, the continued presence of inadequate care is mainly due to difficulties in applying these changes (Shortell et al., 1993; Grol and Grimshaw, 2003). The transformative "Future in Mind" report (Department of Health, 2015) suggested deviating from a generic, health-centred model to a more responsive model for the specific needs of local young populations. It called for service providers and funders to overhaul children and young people's mental health (CYPMH) services into comprehensive systems, offering a range of services from prevention to risk management. This study focuses on reforming a flawed system of service provision, characterised by disunity, inefficiency, and limited access to services.

Since 2016, Child and Adolescent Mental Health Services (CAMHS) within England's National Health Service (NHS) have seen significant changes, advancing more comprehensive approaches to care (Rocks et al., 2020). The National i-THRIVE Programme (NIP) assists CAMHS sites in adopting the THRIVE principles, which prioritise patient needs and cohesive service provision through collaborative care networks (Moore et al., 2023). i-THRIVE was created following implementation science guidelines to facilitate the adoption of THRIVE principles by CAMHS (Moore et al., 2016).

The NIP is explained in a published study protocol (Moore et al., 2023). Briefly, i-THRIVE translates the complex aspects of a 'THRIVE-like' system into practical structures and tools for CAMHS to use in their transformation. The NIP guides CAMHS staff and leaders in developing local models based on THRIVE principles and creating detailed plans for implementation over four phases. Over 65% of children and young people (CYP) live in areas where i-THRIVE has been adopted (i-THRIVE Team). Although in use since 2016, the effectiveness of NIP as an implementation model has not been assessed. This study evaluates the alignment of frontline staff with THRIVE principles, and the effectiveness of NIP in implementing these principles in CAMHS sites.

The THRIVE principles of care encompass characteristics at three levels: macro, meso, and micro. At the macro level, characteristics include interagency function and cooperation. This

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

means a CAMHS system following the "THRIVE-like" approach would involve supporting bodies such as educational and social services in its policy-making and service delivery (Moore et al., 2023). At the meso level, they include a needs-based perspective focusing on CYP and their support services. A site adhering to meso-level THRIVE principles would be expected to have a network of community providers (Moore et al., 2023). At the micro level, these characteristics involve the interactions between CYP, their families, and healthcare professionals (Moore et al., 2023). The embodiment of THRIVE principles depends not only efforts by the CAMHS site but also on broader community involvement, requiring effective working relationships among local systems.

We focus on the average effect of NIP on the implementation sites. For this estimate, we need to know the outcome at each site if NIP had not been implemented (the potential outcome or unobserved counterfactual) (Stuart, 2010). The difference in differences (DiD) approach compares changes in outcomes between implementation and control sites before and after the intervention (Stuart et al., 2014). A key assumption for DiD is that the average outcomes for both groups would have similar trends over time (Xu, 2017), which may not always be plausible. Participation in NIP is voluntary, meaning sites that choose to participate may differ from those that do not. The four-group propensity score-weighted DiD method overcomes these issues by adjusting the implementation and control groups while considering time factors. This adjustment ensures comparability among the groups and reduces bias in estimating the desired effect. It's particularly useful when group composition changes during the study, such as when health practitioners move between Clinical Commissioning Groups (CCGs).

In this study, we assess the effect of the NIP on implementation sites using the four-group propensity-score weighted DiD method. Additionally, we examine how the effectiveness of local system working relationships influences the impact of the NIP.

Methods

Study Setting and Design

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The study protocol is detailed elsewhere (Moore et al., 2023). In brief, we selected twenty CAMHS sites in England: ten of these sites have been using i-THRIVE since 2016 (NIP/implementation sites), and ten others were using different approaches for transformation (comparison sites). Details on the delivery of NIP are in Supplemental Material S1. This study reporting conforms to the Template for Intervention Description and Replication (TIDieR) statement (Hoffmann et al., 2014); the TIDieR checklist is provided in Supplemental Material S2.

The NIP sites were Bexley, Cambridgeshire and Peterborough (CambPeter), Camden, Hertfordshire (Herts), Luton, Manchester (MancSal), Stockport, Tower Hamlets (TowerHam), Waltham Forest (WalthamF), and Warrington. The comparison sites included Bradford, Ipswich and East Suffolk (EastSuff), Lewisham, Norfolk, Northampton (NeneCorby), Portsmouth, Southampton, Stoke-on-Trent (Stoke), Sunderland, and South Worcestershire (Worcester). Additional site details are in Supplemental Material S3.

Surveys

Two surveys were designed based on the RE-AIM Adoption Framework (Glasgow et al., 2019). The staff survey assessed the staff awareness of THRIVE and use of THRIVE principles. Site leaders distributed the survey to professionals involved in providing and commissioning CYPMH services. The survey was administered in August 2019, October 2019, and January 2020.

A survey for transformation leads was also conducted. This survey, completed by programme managers, gathered information about site transformation activities and the support they received. It was distributed from January to May 2020.

Measurements and Data

The primary focus is the degree to which sites follow the THRIVE principles of care, the reporting of which conforms to the Revised Standards for Quality Improvement Reporting Excellence (SQUIRE 2.0) statement (Ogrinc et al., 2016); the SQUIRE checklist is provided in Supplemental Material S4. The THRIVE principles of care encompass macro-, meso-, and micro-level features. We assessed these features in participating sites using the i-THRIVE

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Assessment Tool (Moore et al., 2023). Sites received an overall fidelity and level-specific scores. Higher scores indicate better adherence to THRIVE principles. Fidelity to i-THRIVE was measured at each site before and after the intervention by at least two independent evaluators, with scores averaged for each site and period. When there was not enough information to assign a score, we estimated the missing scores using the mean difference between pre- and post-intervention scores for the site, based on the score for the available period.

To calculate propensity score weights, it is important to identify site characteristics that might cause selection bias or have a confounding effect (Stuart et al., 2014). These characteristics were measured in 2016 and 2019, unless stated otherwise. They included population density (total population per square kilometre) for each CCG (Office for National Statistics, 2016; Office for National Statistics, 2017), annual funding support (£100,000 increments) per CCG (NHS England, 2017a), effectiveness of working relationships in the local system (NHS England, 2021), Indices of Multiple Deprivation (IMD) ranks (Department for Communities and Local Government, 2015; Department for Communities and Local Government, 2019), the initial number of CCGs per site (NHS England, 2019), and compliance in 2017 with the CCG Improvement and Assessment Framework Mental Health Transformation Milestones (NHS England, 2017b). For sites comprising multiple CCGs, we summed the annual funding and averaged the population density, effectiveness of working relationships in the local system, and transformation compliance across CCGs. Site-level IMD ranks were calculated using the method recommended by the Office of National Statistics (Noble et al., 2019).

Statistical Analysis

We analysed the staff survey results using the chi-squared test and Cramer's V as a measure of association. We used a binomial distributed generalised linear model with a log link for site-specific results, excluding "I don't know" responses. To determine if any site had an unusual probability of a "yes" response, we compared the "yes" probability for each site to the average among comparison sites.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Voluntary participation in health policy implementations like the NIP can lead to selection bias, as the sites that choose to participate may differ from those that do not. To correct for this, we applied propensity score weighting (Rosenbaum and Rubin, 1985) to equalise the distribution of characteristics between the implementation and control sites. We used a four-group weighting method (pre-implementation i-THRIVE, post-implementation i-THRIVE, pre-implementation comparison, and post-implementation comparison) to align characteristics across all groups with those of the pre-implementation i-THRIVE sites (Stuart et al., 2014). These propensity scores were calculated using a multinomial model with five site characteristics (population density, annual funding, IMD rank, the number of CCGs per site, and transformation compliance). Characteristic balance was checked using the standardised difference in means (Stuart, 2010). The impact of NIP on fidelity was estimated using maximum-likelihood repeated measures linear regression with an auto-regressive correlation structure, weighted with the calculated propensity scores. To account for remaining characteristic imbalances, we included population density, IMD, and transformation compliance in the final model. The results from this model represent the four-group weighted DiD effect estimate.

To assess the reliability of our fidelity results, we conducted a series of sensitivity analyses. We employed different methods to estimate the effect of the NIP, including the standard (unweighted) DiD and alternative model specifications. We examined the impact of non-compliant control sites by repeating the analysis, excluding these sites. More details can be found in Supplemental Material S5.

To examine variations in the effect of i-THRIVE, we investigated a possible effect moderator, specifically the quality of local system working relationships on i-THRIVE implementation. Further details are available in Supplemental Material S7.

All data visualisation, cleaning, and propensity score modelling was performed in R version 4.2.3 (The R Foundation for Statistical Computing, Vienna, Austria) and RStudio version 2023.03.0 (RStudio, Inc., Boston, MA, USA), with R packages cowplot, extrafont, ggpattern, ggspatial, gridExtra, haven, nnet, RColorBrewer, readxl, reshape2, sf, tableone, and

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

tidyverse (Venables and Ripley, 2002; Wickham, 2007; Neuwirth, 2014; Auguie, 2017; Wickham and Miller, 2017; Pebesma, 2018; Wickham et al., 2019; Wilke, 2020; FC et al., 2022; Yoshida and Bartel, 2022; Chang, 2023; Dunnington, 2023; Pebesma and Bivand, 2023; Wickham and Bryan, 2023). The simple DiD, four-group weighted DiD, and effect modification analyses were performed in SAS® version 3.81 (SAS Institute Inc., Cary, NC, USA) in SAS® Studio.

Results

Surveys

The staff survey had 689 responses across 19 sites (no responses from Luton). Detailed survey results are in Supplemental Material S7.

While the THRIVE Framework was widely known, more implementation respondents recognised it (83.9% vs. 70.5%, $p<0.0001$). A higher proportion of implementation respondents reported using THRIVE principles in their daily practice (58.5% vs. 49.0%, $p=0.03$). Implementation respondents were more likely to score perfectly on a test of the THRIVE Framework (34.1% vs. 22.9%, $p=0.001$).

The transformation leads survey included eight managers from seven implementation sites and eight managers from seven comparison sites. Notably, managers from four comparison sites (Bradford, NeneCorby, Norfolk, and Sunderland) reported using THRIVE as their service transformation model. This prompted a more detailed examination of staff survey results concerning THRIVE implementation at comparison sites.

The site-level analysis of survey results can be found in Supplemental Material S7. Among those who provided a yes/no response at comparison sites, NeneCorby and Norfolk reported a higher likelihood of implementing THRIVE. The odds of respondents reporting site implementation of THRIVE were 4.43 in NeneCorby (95% CI: 1.33–14.80) compared to other comparison sites (76.5% vs. 59.5%), and 4.43 in Norfolk (95% CI: 2.32–8.47) compared to other comparison sites (76.5% vs. 59.5%). Among respondents at comparison sites, there was no difference in personal use of THRIVE principles compared to other

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

comparison sites. Regarding knowledge of the THRIVE Framework, respondents from Norfolk had a higher probability of achieving a perfect score on the quiz: the odds of scoring perfectly were 3.73 (95% CI: 2.25–6.19), compared to other comparison sites (37.7% vs. 22.9%).

THRIVE Fidelity

During the pre-implementation period, i-THRIVE sites had an average fidelity score of 149.0 (range: 132.0—180.2) and comparison sites had an average score of 133.4 (range: 113.0—158.2). Following implementation, i-THRIVE sites had an average score of 166.6 (range: 145.5—195.0), while comparison sites had an average of 142.2 (range: 132.0 —175.0). Two sites had incomplete fidelity score information: the macro-level components for Bexley during the post-implementation period and the meso-level components for Stoke during the pre-implementation period (scores were assigned as outlined in the methods section). Detailed fidelity scores by level and site are illustrated in Figure 1, and a map showing the changes in scores by site throughout the study is presented in Figure 2.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 1: Fidelity Scores by Site

Total fidelity scores during the pre- and post-implementation periods are represented by bar height, with patterned overlay to indicate component levels (macro, meso, micro). Implementation sites are in panel A (blue) while comparison sites are in panel B (red). CambPeter = Cambridgeshire & Peterborough, Herts = Hertfordshire, MancSal = Manchester, TowerHam = Tower Hamlets, WalthamF = Waltham Forest, EastSuff = East Suffolk, NeneCorby = Nene & Corby, SouthHam = Southampton, Stoke = Stoke-on-Trent

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2: Change in Fidelity Scores Over Study Period

The difference in total fidelity scores during the pre- and post-implementation periods are represented by colour, with an increased score in green and a decreased score in red, on a background map of Clinical Commissioning Groups in panel A. Study sites have a bold outline (comparison sites in red, implementation sites in blue). Inset maps for North West/Midlands, London, and South East are in panels B–D, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Site characteristics, adjusted using four-group propensity-score weighting, are in Table S5, Supplemental Material S7. In our analysis of the weighted standardised differences, we identified some remaining imbalances, specifically in population density for the pre-implementation control group, IMD rank for both control groups, and transformation compliance for both control groups. These covariates were included in our effect estimate model, ensuring a more accurate assessment of NIP impact.

The NIP effect estimates are presented in Table 1. Our analysis reveals that the overall fidelity scores were moderately influenced by the NIP. Specifically, i-THRIVE sites showed an average improvement of 7.05 points (95% CI: -4.47–18.57). The most notable improvements were at the macro level, where i-THRIVE sites increased by an average of 2.92 points (95% CI: -1.09–6.92), followed by the meso level with an average increase of 2.76 points (95% CI: -1.98–7.51), and the micro level with an average increase of 1.39 points (95% CI: -3.94–6.72).

When comparing the four-group weighted DiD with the standard DiD analyses (Supplemental Material S5), we found comparable effects on overall and macro-level fidelity. There were shifts in the impacts on meso- and micro-level fidelity. This suggests that lower-level fidelity was more sensitive to the disparities between i-THRIVE and comparison groups, which were corrected through the four-group propensity-score weighting approach. Alternative modelling approaches produced results similar to our analysis. The exclusion of non-compliant comparison sites did not alter the results.

We observed a moderating effect of local systems working relationship effectiveness on the impact of the NIP for overall and macro-level fidelity. i-THRIVE was found to be more effective at sites with highly effective working relationships. The detailed results are in Table 2. i-THRIVE sites with highly effective working relationships showed increased fidelity scores compared to comparison sites with highly effective working relationships. The most significant impact was on overall fidelity scores (16.41, 95% CI: 1.69–31.13), followed by macro-level scores (6.95, 95% CI: 2.15–11.75). The moderating influence of highly-effective working relationships on meso-level and micro-level fidelity was modest (5.52, 95% CI: -

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

0.66–11.71 and 3.95 points, 95% CI: -3.24–11.15, respectively). Notably, there was no discernible impact of the NIP on sites with ineffective working relationships across any fidelity level.

Discussion

A major obstacle in implementing any complex intervention is ensuring it becomes a routine practice across the entire organisation. This concept is often described as making it 'the way we do things around here' (Haines et al., 2004; Proctor et al., 2011). Both THRIVE and i-THRIVE are frequently mentioned in child mental health policy documents, by NHS England's regional transformation boards, and in media related to child mental health. It is encouraging for policymakers to note that over 70% of staff at comparison sites were aware of THRIVE, and nearly 23% exhibited perfect knowledge of THRIVE principles. Despite this high level of awareness, staff at i-THRIVE sites were significantly more likely to be familiar with, understand, and apply THRIVE principles in their daily work. This indicates that the NIP aids in embedding THRIVE principles within an organisation.

When introducing new health policies, understanding their impact and identifying the contexts in which they are most effective is crucial. Estimating the mean effect presents several methodological challenges, especially when data from pre-implementation periods are scarce or when only a few sites are involved. Unlike randomised controlled trials, the adoption of health policies is not random. This means that the characteristics of the implementation and control sites are likely to vary, leading to potential selection bias. Yet, it is a useful methodological approach for evaluating complex interventions selected and implanted by policymakers. It is often unclear which characteristics influence a site's decision to adopt a health policy. Even with numerous characteristics measured, careful consideration of each characteristic's role is necessary. Confounding is another potential bias, where some characteristics may influence both the decision to implement and the site's capacity to do so effectively.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

To adjust for these biases, various methods are available, but these methods can themselves introduce bias. Sensitivity analyses are essential in gaining a deeper understanding of the data. The four-group propensity-score weighting DiD method is designed to mitigate potential selection biases and confounding factors (Stuart et al., 2014). However, a key assumption of DiD, the parallel trends assumption, is not verifiable. Violations of the parallel trends assumption can lead to issues with time-varying confounding (Xu, 2017), complicating the interpretation of results. Thus, while our study provides valuable insights into the effectiveness of the NIP, these issues must be considered when interpreting the findings.

The future direction of this research will involve comprehensive evaluations of the service and clinical impacts of the NIP. These evaluations will encompass a range of critical factors, including the accessibility and efficiency of services, clinical outcomes, patient experiences, and the specific impacts on various sub-groups within the patient population. The latter will particularly focus on racial or ethnic minorities and distinct diagnostic categories, ensuring a broad and inclusive understanding of the NIP's effectiveness.

Conclusions

This study's investigation into the effectiveness of the NIP in England offers significant insights into the implementation of complex health interventions, particularly in the field of CAMHS. The findings underscore the importance of effective working relationships among local systems in the successful adoption and implementation of health policies like i-THRIVE. Specifically, the study demonstrates that sites with highly effective working relationships exhibit substantial improvements in adhering to THRIVE principles, as evidenced by the increase in fidelity scores across macro, meso, and micro levels.

The broad awareness of the THRIVE framework among staff, even in comparison sites, highlights the programme's permeation in the field of child mental health. However, implementation strategies are critical in embedding these principles more deeply within

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

organisations. This distinction is crucial for policymakers and healthcare leaders aiming to foster more effective, integrated mental health services tailored to the needs of CYP. Methodologically, the study navigates the challenges of evaluating health policy implementations in non-randomised settings. The use of four-group propensity-score weighted DiD analysis is an effective approach to address potential biases, such as selection bias and confounding factors. However, it also highlights the inherent complexities and limitations in evaluating policy impact in real-world settings. The study's sensitivity analyses further strengthen the validity of its findings.

In conclusion, the NIP presents a promising model for CAMHS. Its emphasis on effective working relationships and the tailored approach to implementation are key factors in its success. The insights from this study contribute valuable knowledge to the ongoing efforts to improve mental health services and can guide future policies and programmes aimed at enhancing the well-being of CYP.

Financial support

This study is independent research funded by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration North Thames (grant number CRJM, 46AZ74-/RCF/CLAHRC/UCL004, 52AZ95/AA2/UCL5, 52AZ95/AA2/UCL6, UCLP) and UCLPartners, both awarded to AM. The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR, the Department of Health and Social Care or UCLPartners. All research at the Department of Psychiatry in the University of Cambridge is supported by the NIHR Cambridge Biomedical Research Centre (BRC-1215-20014) and NIHR Applied Research Centre. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Conflicts of Interest

AM was the lead for i-THRIVE between January 2015 and January 2017 during which she led the development of the i-THRIVE implementation approach and established the

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

community of practice. She was an employee of one i-THRIVE implementation site during the evaluation.

LE was an employee of one i-THRIVE implementation site during the evaluation.

PF is currently on the board of the i-THRIVE Implementation Programme. Anna Freud Centre hosted the development and on-going implementation of the THRIVE Framework in collaboration with the Tavistock & Portman NHS Foundation Trust. He is Programme Director of Mental Health at UCLPartners which provides programme management of the i-THRIVE Implementation Programme.

Ethical Standards

The NHS/University Joint Research Office reviewed this study and classified it as a service evaluation, meaning it does not need ethical approval or a review by the Research Ethics Committee (IRAS application number: 250439). For the qualitative parts of the study, participants included implementation leads, managers, and front-line staff, who were interviewed about their service transformation projects. At the start of each interview, verbal informed consent was obtained, which included explicit permission for the interview to be recorded, analysed, and used in the service evaluation. This consent was recorded, and transcripts of each interview were made to document it.

The data for the quantitative evaluation is retrospective and was routinely collected from service records. It was anonymised at the source by the business intelligence staff at each trust, with all identifying information removed. After review by local information governance leads, this anonymised data was provided to the evaluation team. As the data is retrospective and de-identified, consent for its use in the evaluation is neither legally nor ethically required. The requirement for ethical approval was therefore waived by the joint research office.

Availability of Data and Materials

All datasets and code are accessible on GitHub (Sippy, 2023).

Acknowledgements

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The authors are indebted to the clinicians, managers and academics who have created THRIVE and their generous help and thoughtful advice throughout the evaluation to assess its effectiveness, in particular Sophie Dunn. We would also like to thank Rudolf Cardinal for help with the research.

References

Auguie B (2017) gridExtra: Miscellaneous functions for “Grid” graphics.

Chang W (2023) extrafont: Tools for Using Fonts.

Department for Communities and Local Government (2015) English Indices of Deprivation 2015 File 5 Scores for the Indices of Deprivation.

Department for Communities and Local Government (2019) English Indices of Deprivation 2019 File 5 Scores for the Indices of Deprivation.

Department of Health (2015) Future in mind: Promoting, protecting and improving our children and young people’s mental health and wellbeing.

Dunnington D (2023) ggspatial: Spatial data framework for ggplot2.

FC M, Davis TL, ggplot2 authors (2022) ggpattern: ’ggplot2 Pattern Geoms.

Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, Ory MG, Estabrooks PA (2019) RE-AIM Planning and Evaluation Framework: Adapting to New Science and Practice With a 20-Year Review. Frontiers in Public Health **7**: 64

Grol R, Grimshaw J (2003) From best evidence to best practice: Effective implementation of change in patients’ care. Lancet **362**: 1225–1230

Haines A, Kuruvilla S, Borchert M (2004) Bridging the implementation gap between knowledge and action for health. Bulletin of the World Health Organization **82**: 724–731; discussion 732

Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, Altman DG, Barbour V, Macdonald H, Johnston M, et al (2014) Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. British Medical Journal Clinical Research **348**: g1687

i-THRIVE Team National i-THRIVE Community of Practice. Anna Freud National Centre for Children and Families

Moore A, Jenkins P, Harris R, Fonagy P, Wolpert M (2016) The crisis in CAMHS: Can i-Thrive provide a solution? British Association for Counselling and Psychotherapy Children, Young People and Families 4–10

Moore A, Lindley Baron-Cohen K, Simes E, Chen S, Fonagy P (2023) A protocol for a multi-site cohort study to evaluate child and adolescent mental health service transformation in England using the i-THRIVE model. PLoS ONE **18**: e0265782

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Neuwirth E (2014) RColorBrewer: ColorBrewer Palettes.

NHS England (2017a) Mental Health Five Year Forward View Dashboard.

NHS England (2021) CCG Improvement and Assessment Framework (CCG IAF) Data Extract 2017/18.

NHS England (2019) CCG Allocations 2019/20 to 2023/24 - Technical Guidance.

NHS England (2017b) Mental Health Five Year Forward View Dashboard.

Noble S, McLennan D, Noble M, Plunkett E, Gutacker N, Silk M, Wright G (2019) The English Indices of Deprivation 2019.

Office for National Statistics (2016) CCG (July 2015) Ultra Generalised Clipped Boundaries in England.

Office for National Statistics (2017) CCG (April 2018) Ultra Generalised Clipped Boundaries in England.

Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D (2016) SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. British Medical Journal Quality & Safety **25**: 986–992

Pebesma E (2018) Simple features for R: Standardized support for spatial vector data. The R Journal **10**: 439–446

Pebesma E, Bivand R (2023) Spatial data science: With applications in R. doi: 10.1201/9780429459016

Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, Griffey R, Hensley M (2011) Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Administration and Policy in Mental Health **38**: 65–76

Rocks S, Fazel M, Tsiachristas A (2020) Impact of transforming mental health services for young people in England on patient access, resource use and health: a quasi-experimental study. British Medical Journal Open **10**: e034067

Rosenbaum PR, Rubin DB (1985) Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score. The American Statistician **39**: 33

Shortell SM, Gillies RR, Anderson DA, Mitchell JB, Morgan KL (1993) Creating organized delivery systems: The barriers and facilitators. Hospital and Health Services Administration **38**: 447–466

Sippy R (2023) rsippy/THRIVE-fidelity: submission. doi: 10.5281/ZENODO.10402515

Stuart EA (2010) Matching Methods for Causal Inference: A Review and a Look Forward. Stat Sci. doi: 10.1214/09-STS313

Stuart EA, Huskamp HA, Duckworth K, Simmons J, Song Z, Chernew ME, Barry CL (2014) Using propensity scores in difference-in-differences models to estimate the

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

effects of a policy change. Health Services and Outcomes Research Methodology **14**: 166–182

Venables WN, Ripley BD (2002) Modern Applied Statistics with S, 4th Edition. Springer, New York

Wickham H (2007) Reshaping Data with the reshape Package. Journal of Statistical Software **21**: 1–20

Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, et al (2019) Welcome to the Tidyverse. Journal of Open Source Software **4**: 1686

Wickham H, Bryan J (2023) readxl: Read Excel Files.

Wickham H, Miller E (2017) haven: Import and Export “SPSS”, “Stata” and “SAS” Files.

Wilke CO (2020) cowplot: Streamlined plot theme and plot annotations for “ggplot2.”

Xu Y (2017) Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. Political Analysis **25**: 57–76

Yoshida K, Bartel A (2022) tableone: Create “table 1” to describe baseline characteristics with or without propensity score weights.

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1: Estimates of Association Between the National i-THRIVE Programme and THRIVE Fidelity

Outcome	Estimated Score Change	95% CI	p-value
Overall Fidelity	7.05	-4.47–18.57	0.212
Macro-level Fidelity	2.92	-1.09–6.92	0.141
Meso-level Fidelity	2.76	-1.98–7.51	0.234
Micro-level Fidelity	1.39	-3.94–6.72	0.588

CI = confidence interval

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

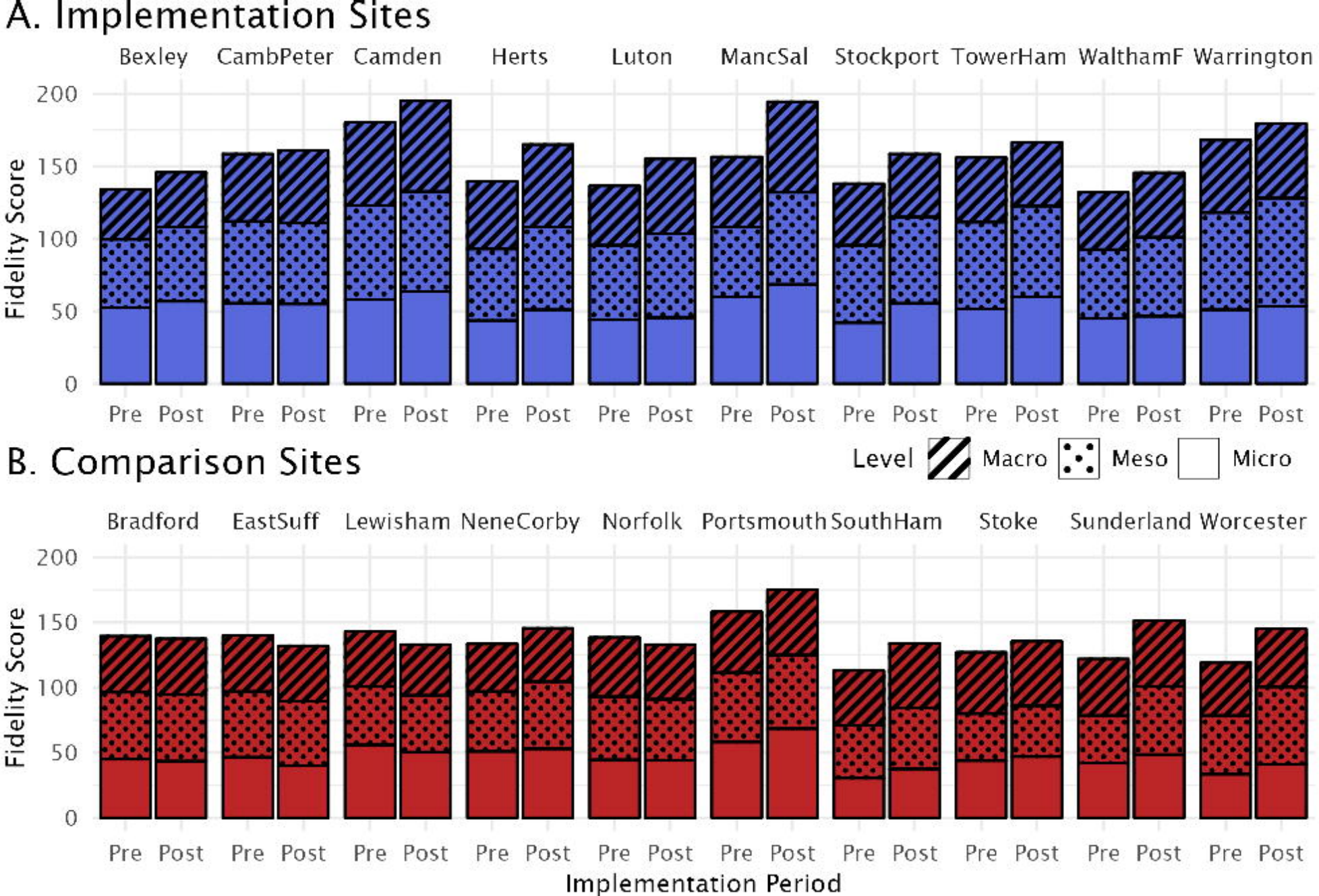
Table 2: Effect Modification by Working Relationship Quality on the Impact of the National i-THRIVE Programme

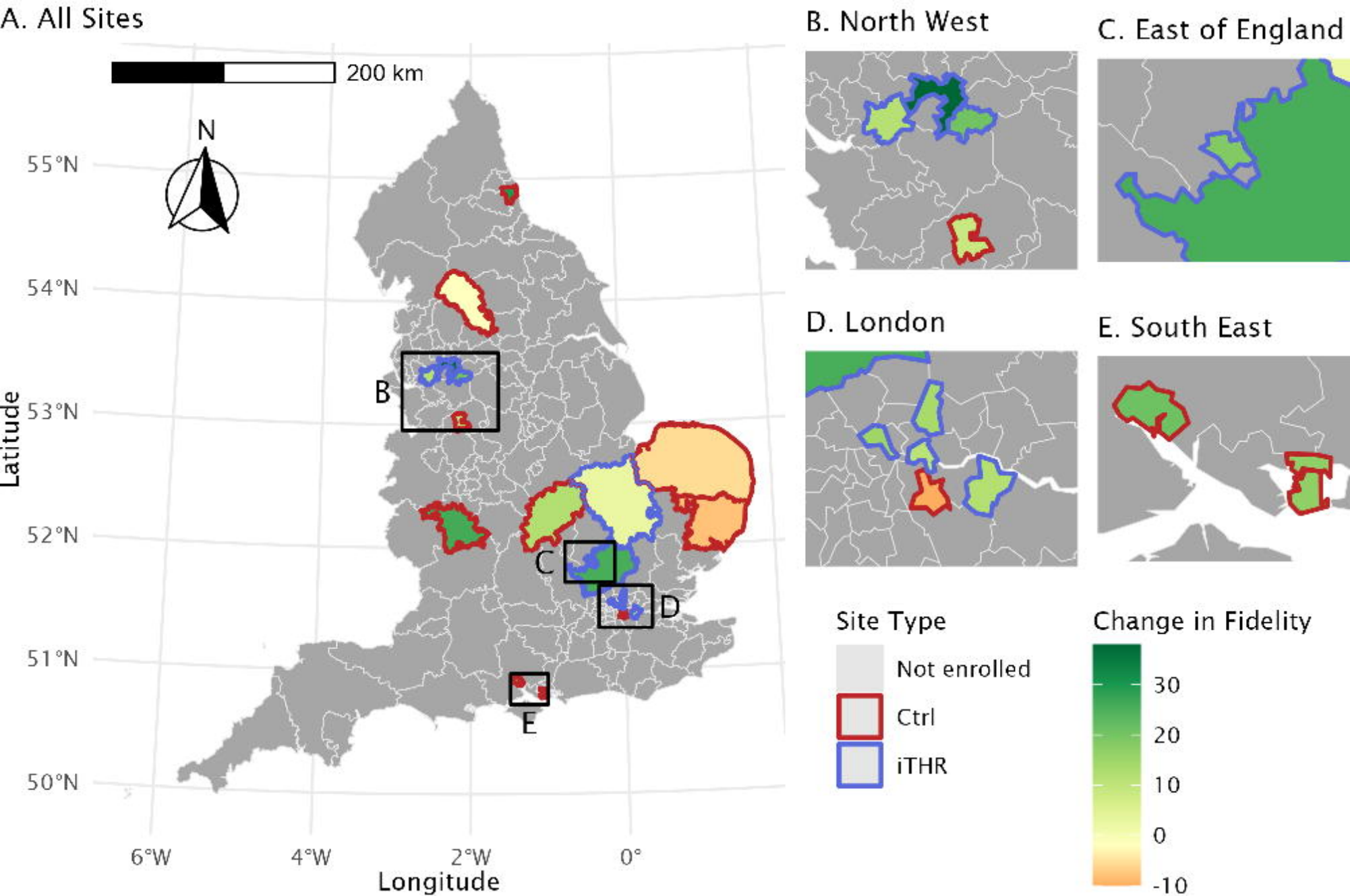
Outcome	Working Relationship Quality	Estimated Score Change	95% CI	p-value
Overall Fidelity	Highly effective	16.41	1.69–31.13	0.031
	Ineffective	-2.76	-18.25–12.73	0.708
Macro-level Fidelity	Highly effective	6.95	2.15–11.75	0.007
	Ineffective	-1.24	-6.29–3.80	0.605
Meso-level Fidelity	Highly effective	5.52	-0.66–11.71	0.076
	Ineffective	-0.75	-7.26–5.76	0.808
Micro-level Fidelity	Highly effective	3.96	-3.24–11.15	0.257
	Ineffective	-0.78	-8.35–6.79	0.828

CI = confidence interval

medRxiv preprint doi: https://doi.org/10.1101/2024.05.07.24306984; this version posted May 7, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

504


A. Implementation Sites
Bexley
CambPeter
Camden
Herts
Luton
MancSal
Stockport
TowerHam
WalthamF
Warrington
Fidelity Score
200
150
100
50
0
Pre
Post
B. Comparison Sites
Level
Macro
Meso
Micro
Bradford
EastSuff
Lewisham
NeneCorby
Norfolk
Portsmouth
SouthHam
Stoke
Sunderland
Worcester
Implementation Period


A. All Sites
200 km
N
55°N
54°N
53°N
52°N
51°N
50°N
Latitude
6°W
4°W
2°W
0°
Longitude
B
C
D
E
B. North West
C. East of England
D. London
E. South East
Site Type
Not enrolled
Ctrl
iTHR
Change in Fidelity
30
20
10
0
-10