Funded by UK Government

TASO Evaluation Report: Analytics for Wellbeing Trials (Northumbria University, Newcastle)

Technical report

March 2026

Project title: Randomised controlled trials to test the impact of wellbeing interventions prompted by learning analytics in higher education (Northumbria University, Newcastle)
Higher education provider: Northumbria University, Newcastle
Evaluators:

- **Impact Evaluation:** The Policy Institute, King's College London
- **Implementation and Process Evaluation:** Northumbria University, Newcastle

Project contributors:

- **The Policy Institute, King's College London:** Dr Susannah Hume, Beti Baraki, Parnika Purwar, Megan Liskey, Professor Michael Sanders
- **Northumbria University, Newcastle:** Dr James Newham, Dr Carly Foster
- **TASO:** Dr Rob Summers, Christoph Koerbitz, Luke Arundel, Mikayla Boginsky
- **QA (Impact Evaluation):** Dr Patrick Taylor (BIT)

TASO

King's College London

Northumbria University NEWCASTLE

TASO Evaluation Report: Analytics for Wellbeing Trials (Northumbria University, Newcastle)

Technical Report

Project title: Randomised controlled trials to test the impact of wellbeing interventions prompted by learning analytics in higher education (Northumbria University, Newcastle)

Higher education provider(s): Northumbria University, Newcastle

Evaluator:

- Impact Evaluation: The Policy Institute, King's College London
- Implementation and Process Evaluation: Northumbria University, Newcastle

Project contributors:

- **The Policy Institute, King's College London:** Dr Susannah Hume, Beti Baraki, Parnika Purwar, Megan Liskey, Professor Michael Sanders
- **Northumbria University, Newcastle:** Dr James Newham, Dr Carly Foster
- **TASO:** Dr Rob Summers, Christoph Koerbitz, Luke Arundel, Mikayla Boginsky
- **QA (Impact Evaluation – v1.0):** Dr Patrick Taylor (BIT)

VERSION	DATE	REASON FOR REVISION/NOTES
1.1	10 March 2026	Updated introduction to include more recent citations. Copy edited report to fix grammatical errors and improve readability.
1.0 [original]	7 January 2026	Reviewed by external reviewer.
Pre-registration	29 October 2024	This design has been pre-registered on Open Science Framework[1]

[1] TASO (2024, October 29). Analytics for Wellbeing trials – Northumbria University. Retrieved from osf.io/kj4sb

Contents

Acknowledgements

This project was funded by a grant from the Evaluation Task Force (Cabinet Office/HM Treasury) and its Evaluation Accelerator Fund (2023-2025).

1. Summary

This study evaluates the impact of a wellbeing-focused learning analytics intervention delivered at Northumbria University (NU) to support undergraduate students exhibiting early signs of academic disengagement.

This evaluation examines whether it is more effective to identify students for wellbeing support using their WHO-5[2] wellbeing score, collected at enrolment, or their learning analytics risk rating (LAR), [3] generated by NU's learning analytics system.

Table 1: Summary of evaluation details

Intervention name	Targeted wellbeing interventions prompted by analytics [4]
Intervention type	Learning analytics (post-entry)
Higher education provider	Northumbria University
Evaluator	The Policy Institute, King's College London (Impact Evaluation) Northumbria University (Implementation and Process Evaluation (IPE))
Evaluation design	Randomised controlled trial (RCT) IPE Trial participants were randomised into treatment groups and a control group: • $TG1_{WHO-5:}$ participants are identified to receive 'nudge' emails based on their WHO-5 score being below 50 at enrolment.

[2] The WHO-5 is a self-report instrument measuring mental wellbeing. It consists of five statements relating to the past two weeks. Each statement is rated on a 6-point scale, with higher scores indicating better mental wellbeing.

[3] LAR refers to the learning analytics generated probability of a positive educational outcome for a student.

[4] The title of the intervention has been updated from "Wellbeing interventions prompted by analytics" as indicated in the trial protocol.

	• $TG2_{LAR}$: participants are identified to receive 'nudge' emails based on their Learner Analytics Risk Rating (LAR) dropping below 75.
Comparison group type	Two treatment groups (TGs), where nudges are either prompted by score on a wellbeing survey ($TG1_{WHO-5}$) or by analytics ($TG2_{LAR}$) One control group that receive no nudges
Sample size	13,122 students • 4,367 students in $TG1_{WHO-5}$ • 4,367 students in $TG2_{LAR}$ • 4,388 students in the control group
Outcome measures	Primary outcomes: • Whether student referred themselves to support services Secondary outcomes: • Whether student self-referred to Counselling and Mental Health Team (CMH) • Whether student self-referred to SilverCloud (online guided self-help) • Email 'click-through' rate (between treatments only) to any of the links in the intervention email • LAR (continuation prediction score) at the start of the following semester • Withdrawal from university

The Policy Institute at King's College London (KCL) and NU designed a three-arm randomised controlled trial (RCT), which was delivered by NU, to assess whether using risk metrics, such as WHO-5 wellbeing scores at enrolment or learning analytics risk rating (LAR), can more effectively target wellbeing support for students.

Students randomised to the treatment groups, who were identified as at risk on either metric, were further randomised to receive one of two targeted emails. Students in the treatment groups, who were not identified as at risk, received a generic email. The control group did not receive any emails.

The RCT reported in this impact evaluation examined whether different methods of identifying at-risk students, who then received targeted emails or no communication, resulted in higher engagement with support services, with the emails themselves, and with the university.

In addition, NU designed and delivered IPE of this learning analytics prompted email-based intervention. Using a mixed-methods approach, the IPE focused on how the intervention was delivered and explored the experiences of the staff involved. Drawing on a range of data sources, including a focus group with staff, and administrative datasets, it examined the fidelity of the implementation of the intervention, whether the intervention reached the intended students, differential engagement patterns among students, and if the nudge messaging to students raised the likelihood that students engage. Overall, the IPE complemented the impact evaluation and provided further insights into the factors shaping the intervention's effectiveness.

The RCT included 13,122 undergraduate and postgraduate students who were randomly assigned to one of three groups to evaluate how different methods of identifying at-risk students affected their likelihood of accessing support services. This included the following groups:

- Students in Treatment Group 1 ($TG1_{WHO-5}$) were identified as at risk based on WHO-5 score being less than 50. This group comprised of 4,367 students.
- Students in Treatment Group 2 ($TG2_{LAR}$) were identified as at risk based on LAR score, being less than 75 (representing a <75% likelihood of continuation). This group comprised of 4,367 students.
- Students in Control Group received no contact from the university related to the study, regardless of whether they were identified as at-risk on either metric. This group comprised of 4,388 students.

Students identified as at risk in $TG1_{WHO-5}$ and $TG2_{LAR}$ were subsequently randomised again to receive one of two 'nudge' emails: one signposting to a 'self-help' digital platform called SilverCloud and one signposting to one-to-one support (counselling). Students within these groups who were not identified as at risk on the treatment metric received a generic email. The control group were not randomised again and did not receive any nudge email. This nested RCT, i.e. the second randomisation stage, conducted by NU, fell outside the scope of KCL's evaluation and high-level statistics are reported separately as part of NU's implementation and process evaluation (IPE).

Summary of findings

- There was no significant impact of different methods of identifying at-risk students on students' likelihood of self-referring to support services.
- There was no significant impact of the treatment on the likelihood of self-referring to Counselling and Mental Health (CMH) services.
- There was no statistically significant difference between $TG1_{WHO\text{-}5}$ and $TG2_{LAR}$ and the control group with respect to self-referral to SilverCloud.
- There was no statistically significant impact of the treatment on the other secondary outcomes of interest (LAR, likelihood of clicking on links embedded in the intervention email, and likelihood of withdrawing from university).

The IPE also provided important context to interpret the above results:

- The trial used a complex opt-in consent process, which limited the reach of the intervention. In the end, the trial only included just over half of all eligible students, and consent rates were associated with different demographic characteristics; for example, female students were more likely than male students to consent to the use of their wellbeing survey scores.
- There was little overlap in the population of students identified as having low wellbeing by an analytics system and those identified as having poor wellbeing by a wellbeing survey at enrolment.
- Overall engagement with the nudges was low, only 17.7% of nudges were opened by students and only 2.1% clicked through to the embedded support links.
- Students in $TG1_{WHO\text{-}5}$ were more likely to open and click the emails than those in $TG2_{LAR.}$
- In terms of implementation issues, participants noted the nudges were not timed well as they were sent around half-term and staff cover was limited.
- Staff appreciated the intervention and found it potentially useful for students, but acknowledge there are ongoing capacity issues and would not be able to meet the needs of students without increased resources.

2. Introduction

Poor mental health not only affects students' wellbeing but is also linked to a range of adverse academic and life outcomes. These include reduced educational attainment, disengagement from university, increased dropout rates, and slower progression into employment. Additionally, poor mental health poses significant risks to physical health, including increased incidences of self-harm and suicide (Newham and Francis, 2018).

In recent years there has been growing concern around student mental health and wellbeing, which has intensified following the COVID-19 pandemic. Evidence suggests that the students in higher education (HE) face heightened and multifaceted pressures with various interconnected elements at play (Jones and Bell, 2024). This has been attributed to several factors such as academic workload, isolation and loneliness, financial difficulties and the challenges of adapting to a new environment (Worsley et al., 2020). In addition, a concerning trend is also the disproportionate impact of mental health issues on specific student groups. For example, students from ethnic minority groups, care-experienced students, as well as those from low socio-economic backgrounds are more vulnerable to mental health challenges (Robertson et al., 2022).

Despite the magnitude of these issues, the proportion of HE students who access help for mental health difficulties is much lower than the expected prevalence (McLafferty et al., 2024; Newham and Foster, 2025) However, HE CMH teams are already struggling to meet increasing demand despite increases in funding (Morrish, 2023; Pollard et al., 2021). Consequently, HE institutions are in the precarious situation of needing to empower struggling (and potentially the most at-risk) students to access support while maintaining support to an increasing caseload of students (Birch et al., 2026; Clegg et al., 2026).

With the increasing use of learning analytics to improve student retention and progression, policy has encouraged HE institutions to align these systems with the mental health and wellbeing agenda to identify vulnerable students, target interventions, and provide clear pathways for accessing support (Peck et al., 2025; Department for Education, 2024). By using data for profiling and promotion of specific pathways of support, it may lead to better allocation of scarce resources such as one-to-one counselling and guided self-help, and thus alleviate pressure and service demand (Foster, 2026). Examples of universities attempting to target and personalise student mental health support on a large-scale are starting to emerge (Keane, 2024), but there remains limited published research on their effectiveness in improving outcomes. Furthermore, the translation of testing wellbeing analytics into standard practice, and the impact of introducing such data systems on the practice and roles of those in CMH teams remains unclear (Newham and Foster, 2025). A recommendation of the Higher Education Mental Health Implementation Taskforce

(2024) is to gather views from providers on their understanding of wellbeing analytics, the benefits it might bring, and the institutional changes needed to implement and leverage the benefits of wellbeing analytics.

To investigate this, TASO has commissioned KCL to evaluate the impact of a wellbeing intervention driven by learning analytics at NU. This project aimed to test the efficacy of wellbeing measures and learning analytics to identify students at risk, whilst the NU internal evaluation focussed on randomising communications for different types of support to better understand what factors influence help-seeking behaviour in students.

The remainder of the report is structured as follows: Section 2 describes the intervention's design, objectives, and underlying Theory of Change. Section 3 outlines the impact evaluation delivered by KCL, including research questions, methodology, and findings. Section 4 presents the IPE by NU, detailing how the intervention was delivered and received. Section 5 discusses the implications of the finding for policy and future research. Section 6 outlines the roles and responsibilities of the project delivery and evaluation teams, and Section 7 summarises the ethical considerations and procedure that underpinned the study. Finally, Section 8 and 9 outline the references and appendices respectively.

3. Intervention description

The trial evaluates two different identification strategies as part of which students were randomised into treatment and control groups. In a first stage, at-risk students within the treatment groups were identified based on either the LAR score or the WHO-5 survey score. In a second stage, students identified as at risk on either metric were subsequently randomised to receive one of two targeted messages. Students who were not identified as at risk received a generic email. The intervention is used as a basis for exploring whether different methods of identification and messaging have an effect on student's self-referral behaviour to the services available on campus. The intervention is delivered within NU's student population at its Newcastle and London campuses. Table 2 below outlines the key features of the intervention.

Table 2: Summary of intervention details

Section	Description
Why is the intervention being run?	There is a hidden or undisclosed student population who are not naturally self-help seekers at a time of need and would benefit from receiving personalised email-based nudges to encourage them to access wellbeing support services. The aim is to understand if analytics can be used to identify these students and whether or not the nudges prompted by analytics are effective in driving students to support services.
Who is the intervention for?	Undergraduate students who provided consent for their learning analytics data to be used, and who met the definition of being at risk, using either their WHO-5 score (<50) at enrolment or their LAR score (<75% at Week 4 of Semester 1 of the 2024/25 academic year).

Section	Description
What is the intervention?	Students in the treatment groups, identified as at risk, either via the WHO-5 score or the LAR score, received one of two targeted emails directing them to support services. Students in the treatment groups who were not identified as at risk received a generic email, which signposted wellbeing support available on the student portal. The trial is designed to test the effectiveness of two distinct components: • Comparing two student-at-risk identification methods: the WHO-5 score versus the LAR score. This was tested by KCL. • Comparing two 'nudge' emails: one email signposting one-to-one support (counselling) versus one email that directed students toward online self-help (SilverCloud). This was explored in the IPE by NU.
Who is delivering the intervention?	The delivery team at NU were responsible for sending the emails to students in $TG1_{WHO-5}$ and $TG2_{LAR}$.
How is the intervention delivered?	The emails were delivered via email using the customer relationship management (CRM) system that is used to send marketing email communications to students.
How many times is the intervention delivered?	Once, during Teaching Week 4 (October 2024)

3.1 Theory of Change

A Theory of Change (ToC) is a theoretical model that outlines how an intervention is expected to cause or contribute to a change in outcomes. It explains the logical sequence of the intervention, from its inputs, activities and outputs to proximate and distal outcomes and impacts. A ToC aims to articulate and illustrate the mechanisms and assumptions that explain why an intervention is believed to lead to its hypothesised outcomes.

3.1.1 Theory of Change development

To develop the ToC for NU's analytics-prompted, email-based wellbeing intervention, KCL initially created a draft ToC based on a set of scoping activities and then led a half-day online workshop with representatives from NU to review and develop further its details.

The scoping activities included performing a desk review of intervention-related documentation, which were supplemented by some initial discussions about the intervention with NU as part of the project's inception. These activities helped to develop an initial understanding of the intervention's rationale, and its expected outcomes and impacts. The workshop consisted of group discussions, where attendees outlined and reviewed the intervention's expected inputs, activities, outcomes and impacts, as well as the critical assumptions underlining the relationships between each of these factors. Following the workshop, KCL refined the ToC and shared it with the workshop participants before it was finalised.

3.1.2 Description of Theory of Change

A written summary of the ToC developed for NU's intervention is provided below. An illustration developed by KCL is also included in Appendix B using TASO's core ToC template (TASO, n.d.).

Problem statement: HE students are disclosing more needs in relation to their mental health, and while some proactively access support, there is a hidden or undisclosed student population that do not tend to actively look for support at a time of need. This lack of awareness or engagement in accessing support services can negatively affect students' academic journey and overall wellbeing.

Aim: NU aims to use the analytics systems to reduce complexity in both identifying students who may need wellbeing support and helping them access support services through proactive and tailored signposting that are appropriate to the students' level of need.

Inputs:

- Consent-based platform for collecting wellbeing (WHO-5) data
- Access to analytics insights by student, using the learning analytics system, Hobson CRM student support data, and enrolment data
- Hobson CRM platform
- Staff, including CMH team and analysts' time and knowledge
- Budget
- Data Protection Impact Assessment

Activities:

- Data collection (including consent and ethics) and analysis
- Email-based nudges model development and monitoring (setting up system)
- Personalised email-based nudges – based on analytics (LAR) or WHO-5 (wellbeing) – sent to students offering support packages
- Engagement tracking, including creation and analysis of a master dataset

Outputs:

- Number of students receiving emails based on analytics and wellbeing data
- Number of email-based nudges sent per student
- Number of students receiving email-based nudges who engage with support
- Core10 data
- GDPR compliant dataset

Outcomes:

There are two primary stakeholders: students and the institution of NU, particularly staff who are involved with supporting student engagement. The outcomes are divided into short, medium, and long-term outcomes below, for each of those stakeholders.

Short-term outcomes for students:

- Improved awareness and knowledge of support services and accessibility routes
- Removal of (pre-)contemplation stage among students that are non-self-help seekers

Medium-term outcomes for students:

- Increase students' engagement with appropriate support
- Improved likelihood of continuation

Medium/long-term outcomes for NU:

- Improved ability to identify at-risk students in need of support
- Improved understanding of how to engage students in the support process

Impacts:

- An evidence-based roadmap for implementing a 'whole university' approach to mental health and wellbeing whereby all students receive better guidance on seeking support
- Students re-engage with their studies

3.1.3 Anticipated mechanisms and causal pathways

The following statements further characterise the key causal pathways that explain why it is expected the wellbeing workshops will lead to changes in the hypothesised impacts and outcomes.

Causal pathway 1: Identification and personalised messages

It is anticipated that using analytics (LAR) and wellbeing data (WHO-5) to identify students who may not have come forward to seek support will help NU reach those in need at an earlier stage. Following identification, sending them personalised emails that focus on their current behaviour can increase the likelihood that the messages will be perceived as relevant and timely. This will in its turn increase the likelihood that the emails will be noticed and read by students, compared to the generic emails they regularly receive from the university. In a final turn, this may increase the likelihood that students will actively engage with the emails.

Causal pathway 2: Increased sense of belonging

It is anticipated that students receiving these messages will feel the university is both noticing their need and intervening on their behalf early on. If the students felt they were struggling alone, receiving these messages might give them a sense of relief that the university cares about their wellbeing. This is likely to increase their motivation to engage with the support being offered, particularly among students who felt isolated and disconnected from the university.

Causal pathway 3: Increased awareness of support services

The emails include a clear tailored link to appropriate support services (e.g. online self-help guide or counselling). This removes one of the key barriers for students who may not know what support exists and how to access it. Increased awareness of existing

services increases the likelihood that students will take the first step toward help-seeking behaviour.

Causal pathway 4: Moment for reflection and agency

Receiving tailored and personalised emails with an appropriate support service might give students a moment to reflect on their situation. This moment of reflection can trigger students to change their behaviour to seek support and eventually re-engage with their studies.

Causal pathway 5: 'Whole university' approach adopted

By identifying students in need early-on and signposting them to appropriate support services before students hit crisis points in relation to their mental health or academic performance, NU may be able to develop a system that understands patterns of disengagement and unmet needs of its students. This can inform future interventions and support the development of a 'whole university' approach to meeting student mental health and wellbeing needs.

3.1.4 Discussion of moderating factors

The impact of these pathways may be influenced by several moderating factors including:

- The accuracy of WHO-5 and the LAR in identifying at-risk students who are responsive to prompts to engage with wellbeing support services.
- Whether students open, understand, and trust the emails.
- How the students perceive the tailored and personalised emails. For example, whether they perceive the emails as supportive or as intrusive and punitive.
- The capacity and responsiveness of support services once students seek help.
- The underlying reasons for disengagement, which may require support outside of what the university can provide. For example, if disengagement is due to financial hardship or caring responsibilities, students may need a different type of support that the university is not equipped to provide promptly.

4. Impact evaluation

A key goal of the evaluation was to assess the effectiveness of two different mechanisms for identifying students with poor wellbeing. This was tested through an RCT.

4.1 Research questions

The research questions for the impact evaluation were:

1. To what extent does identifying students based on their WHO-5 score or their LAR effectively identify students who are responsive to prompts to engage with wellbeing support services?
2. To what extent are students targeted based on their WHO-5 score or their LAR more likely to click-through to links in nudge emails signposting them to services?
3. To what extent does identifying students based on their WHO-5 score or their LAR and sending nudge emails signposting them to services improve their LAR at the beginning of the following semester?
4. To what extent does identifying students based on their WHO-5 score or their LAR and sending nudge emails signposting them to services reduce their likelihood of withdrawing at the beginning of the following semester?

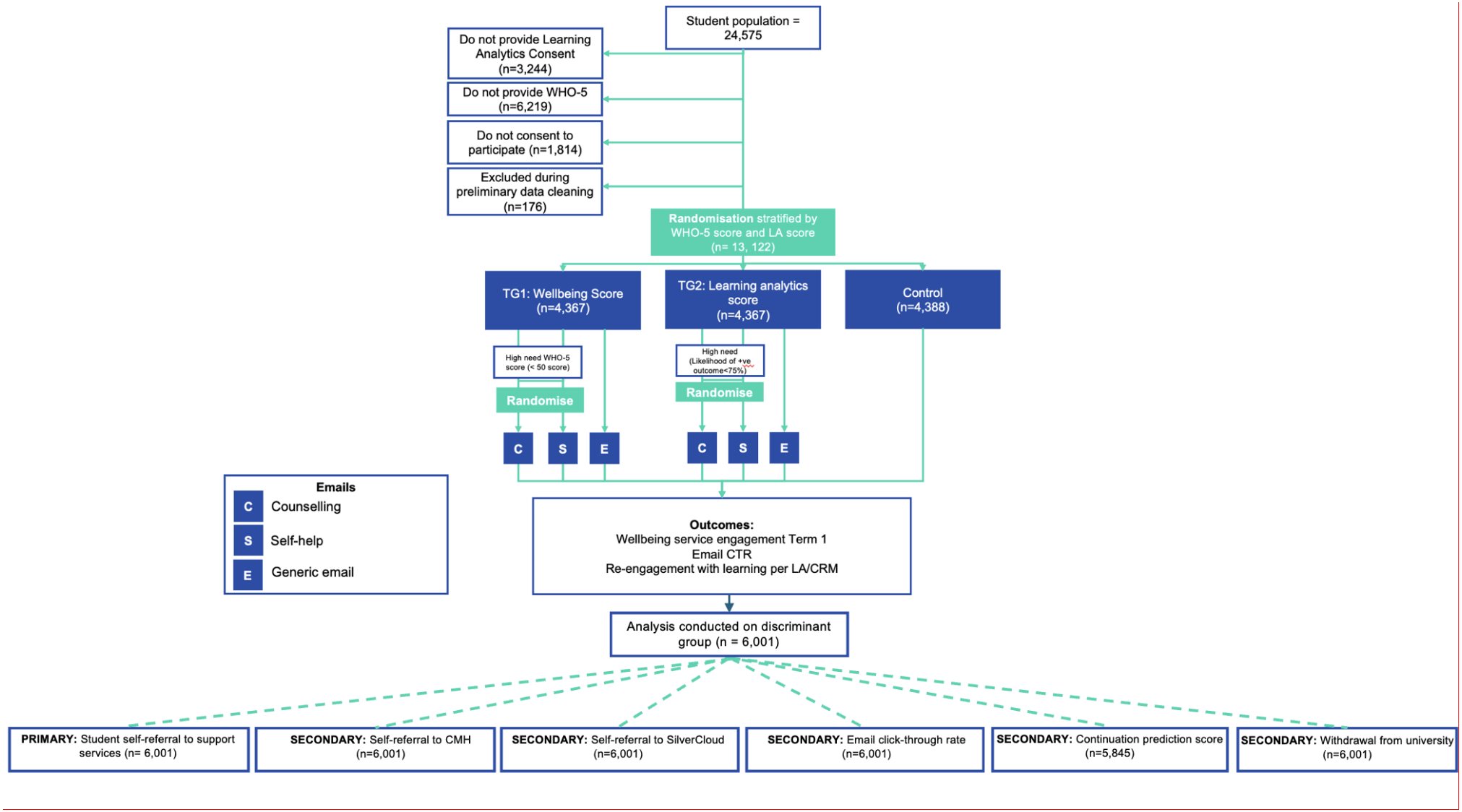
4.2 Impact evaluation design

4.2.1 Design

The impact evaluation was run as a three-armed individually randomised RCT to understand whether identifying students based on their WHO-5 wellbeing score at enrolment or their LAR is more effective for targeting communications on available wellbeing support. The study period ran from September 2024 to January 2025, with final reporting in June 2025.

Figure 1 gives an overview of the study flow up to the point of final data collection.

Figure 1: CONSORT flow diagram of RCT



4.2.2 Sample selection

The study took place at NU, which has three campuses (Newcastle, London, Amsterdam) and distance learning provision. Of these, the Newcastle campus has the majority of students (78%).

Students based at the Newcastle and London campuses, and distance learners based in the UK were all in-scope to be included in the trial. Those who provided consent to the trial were asked to:

- Complete the WHO-5; and
- Provide consent for the learning analytics system.

In the trial, the starting cohort was 24,575 students, of whom:

- 3,244 students did not consent to the use of their educational analytics data
- 6,219 students did not consent to the use of their WHO-5 wellbeing data
- 1,814 students did not provide consent to be part of the trial.

After accounting for the above exclusions and conducting preliminary data cleaning, the final sample comprised of 13,122 students. This aligns with the consent rate of approximately 60% of the total student population that was assumed during the trial design phase and indicated in the trial protocol. After randomisation, there were 4,367 in $TG1_{WHO-5,}$ 4,367 in $TG2_{LAR}$ and 4,388 in the control group.

4.2.3 Randomisation

Randomisation was carried out at the individual level. It was conducted by NU using R code developed by KCL. It was stratified by a median split of WHO-5 and LAR score, creating four strata: those in the top half of risk for just WHO-5 score, those in the top half of risk for just learning analytics score, those in the top half for both scores, and those in the top half for neither score.

Randomisation of the nested study, which is out of scope for the impact evaluation, was carried out by NU. KCL provided R code, which NU adapted.

See Appendix G for both randomisation codes.

4.2.4 Defining the analytical dataset

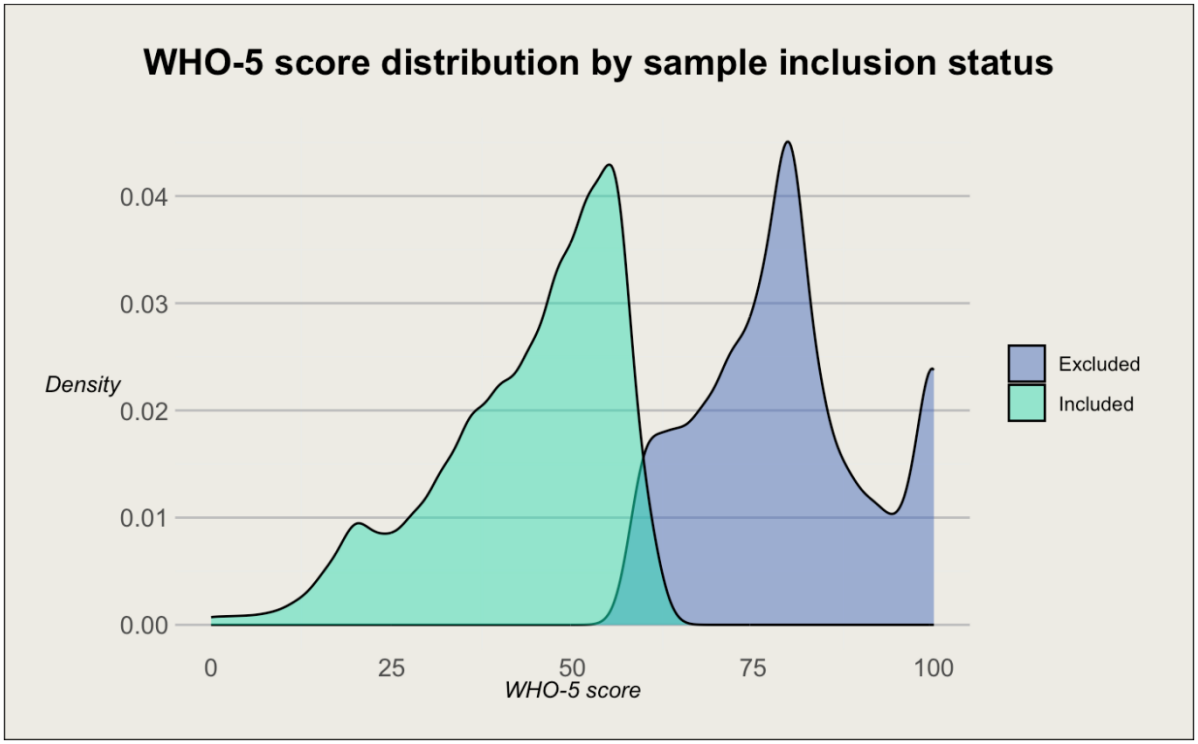
The study faced a trade-off between sample size and the proportion of the sample for analysis that was treated, which if held constant, would cause attenuation of the treatment effect. Because power is lost logarithmically and attenuation occurs approximately linearly, this means that a smaller sample of whom more participants are treated yields more statistical power. This, in turn, must also be traded off for the need

for the intervention to be discriminant in both groups – that is, not everyone assigned to each group at random should be treated, or it becomes impossible to measure the effectiveness of the method of identification.

As a result of the aforementioned trade-off, in this trial, while all the 13,122 participants were randomised, the analysis of primary and secondary outcomes was only conducted on a discriminant group (see Section 5.3.5 for ex ante and ex post power calculations). This sample was created by starting off by sampling the highest risk 3,000 participants based on WHO-5 score, and the highest risk 3,000 participants based on their learning analytics score. This sample was then expanded in an iterative fashion, and 6,001 participants were identified.[5] Of these 6,001, 2,584 were in scope using WHO-5 only, the same number were in scope using LAR only, and 833 were in scope for both risk measures.

Figure 2 and

Figure 3 present the distribution of WHO-5 and LAR scores for the students who were in and out of the scope of the sample using the approach outlined above.



[5] Per the protocol, it was calculated that 6,000 participants would optimise this sample size/proportion treated trade-off.

Figure 2: WHO-5 score distribution by sample inclusion status

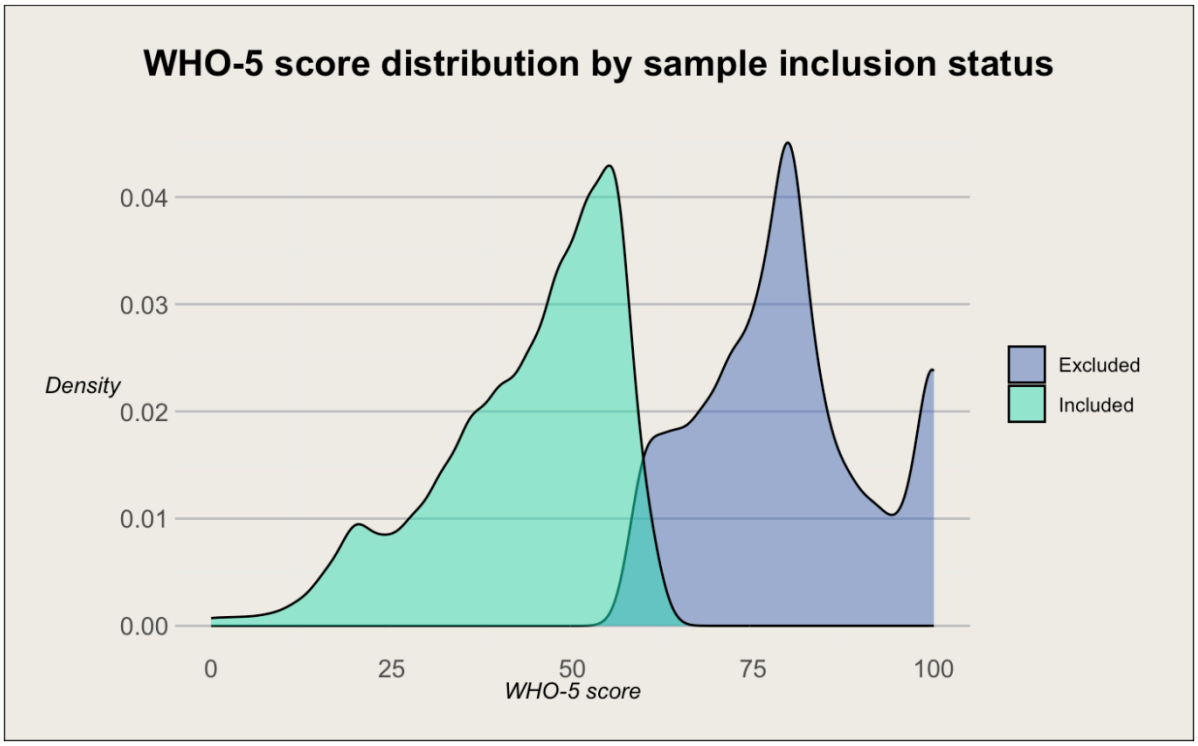
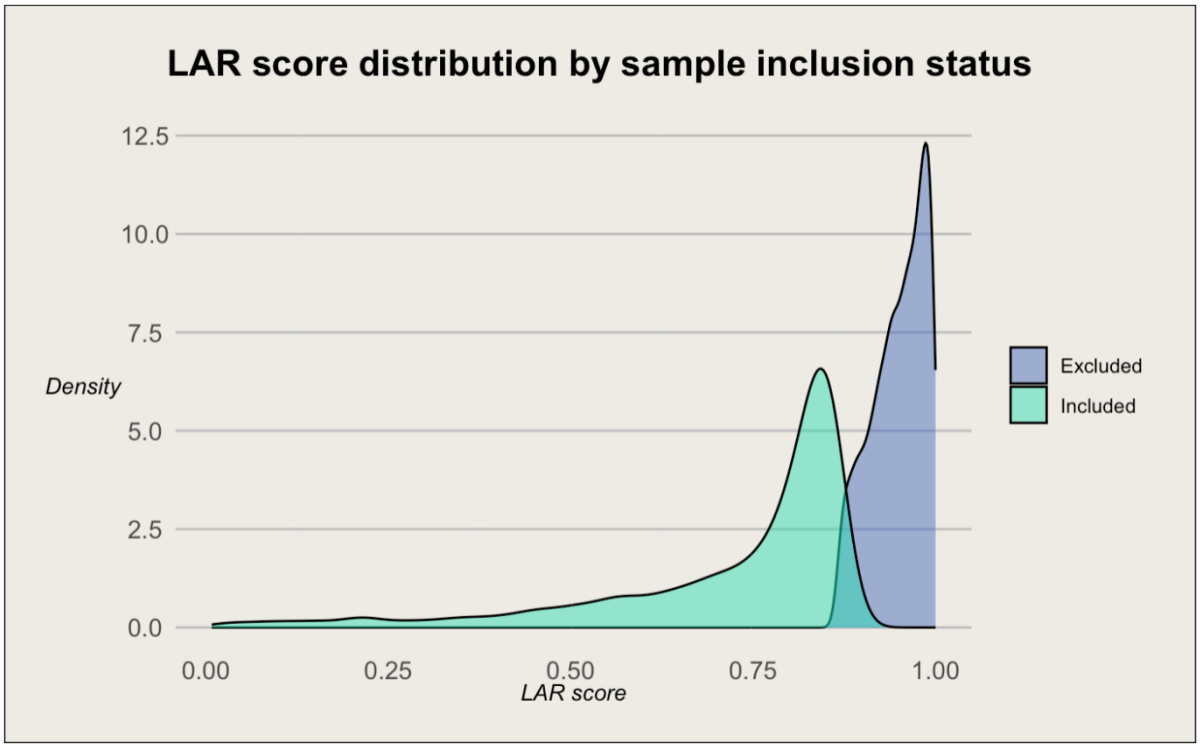


Figure 3: LAR score distribution by sample inclusion status



4.3 Data

4.3.1 Outcome variables

The outcome measures reported in Table 3, below, were chosen in collaboration with NU and represent how the university conceptualises student engagement. Each outcome measure is connected to the data sources pragmatically available for the purpose of the trial to measure said outcome.

Table 3: Outcome measures

Outcome type	Measure	Data source	Point of collection
PRIMARY: Student self-referral to support services	Whether student self-referred to either CMH or SilverCloud **binary**	CRM	End of Semester 1, December 2024
SECONDARY: Self-referral to CMH	Whether student self-referred to CMH **binary**	CRM	End of Semester 1, December 2024
SECONDARY: Self-referral to SilverCloud	Whether student self-referred to SilverCloud **binary**	CRM	End of Semester 1, December 2024
SECONDARY: Email 'Click-through' rate	Whether student clicked through to any of the links in the intervention email **binary** (only available for treatment groups)	Hobsons CRM[6]	November 2024
SECONDARY: Continuation prediction score	Learning analytics generated probability of a positive educational outcome for student **continuous**, 0 - 1	NU Learning Analytics System[7]	Start of Semester 2, January 2025
SECONDARY: Withdrawal from university	Whether student withdrew from NU prior to commencement of Semester 2 **binary**	Data extract downloaded from NU SITS[8] system	Start of Semester 2, January 2025

[6] This is an email communications platform used by NU to send out the nudge emails. It enabled the creation of a bespoke participant list for the purpose of the trial post randomisation as this was not possible via the standard system.
[7] This is a Software-as-a-Service (SaaS) platform licensed by NU. It provides student-level risk scores/ percentages, which can be downloaded and used to support targeted interventions.
[8] This is a students' record system at NU, which captures student information such as demographic characteristics and consent preferences.

4.3.2 Other variables

Table 4 provides the covariates used for sample stratification and in the primary specification for the analysis. All covariates were collected by NU via their student record systems (except LAR, which was collected via the learning analytics system). It was anticipated that not all covariates would be available for analysis for each participant, due to anonymisation requirements. The provision of these data was ultimately up to NU's discretion, balancing the amount of information provided with this duty to maintain anonymity. See Section 4.5 for discussion of why campus, as a covariate, was ultimately excluded.

Table 4: Covariates ranked in order of importance

Data Point	Purpose	Data to be collected	Provided?
WHO-5 baseline score	Stratification variable	0, 4, 8 [...] 100	Yes
LAR (baseline)	Stratification variable	0 – 1	Yes
Campus	Predictor of outcomes, included in model	Newcastle or Not Newcastle	No
Subject	Predictor of outcomes, included in model	CAH1 level (or higher level of aggregation e.g. Faculty)	Yes – as Faculty "Art, Design and Social Sciences (ADSS)", "Business and Law (BL)", "Electrical and Electronic Engineering (EE), "Health and Life Sciences (HLS)"
Gender	Predictor of outcomes, included in model	"M", "F", "Non-binary", "Not disclosed"	Yes

Data Point	Purpose	Data to be collected	Provided?
Ethnicity	Predictor of outcomes, included in model	"White", "BAME", "Other"	Yes – as "White", "Non-White", "Unknown"
Age grouping	Predictor of outcomes, included in model	High level grouping – Factor	Yes – as "Over 21", "21 or younger"
Student status (new or continuing)	Predictor of outcomes, included in model	"New", "Continuing"	Yes
First in family to attend university	Predictor of outcomes, included in model	"Yes", "No"	Yes – with addition of "Unknown"
Level of study (UG and PG)	Predictor of outcomes, included in model	"UG", "PG"	No
International student fee status	Predictor of outcomes, included in model	"Home", "International"	No

4.3.3 Data collection methods

Data were collected by NU via their management information systems and were shared with KCL for analysis. Datasets were linked by an anonymous key held by NU. Attrition from outcomes was less than 5% as outcomes were collected via NU management information systems.

4.4 Analytical strategy

4.4.1 Descriptive analysis

The descriptive analysis has been presented to provide an overview of the key variables across the full sample, as well as separately for the treatment and control groups. This included analysis of key socio-demographic characteristics such as gender, ethnicity, mature student status, IMD quintile, study year, disability status and school of study.

4.4.2 Balance Checks

The following rules were used to conclude if a covariate was imbalanced in terms of its distribution across the treatment and control group:

- For continuous variables: if the absolute difference in the means between the two groups, as a proportion of the sample standard deviation (equivalent to a Z-score within a Standard Normal Distribution), exceeded 0.1.
- For binary variables: if there was a difference of more than five percentage points in the proportions of respondents in each category.

4.4.3 Analysis of primary and secondary outcomes

Regression analysis was conducted to examine the impact of the mechanism of identifying students (WHO-5 or LAR) to be sent one of two 'nudge' emails on the outcomes specified in Table 3.

The analysis estimates the impact of being assigned to receive an augmented 'nudge' email based on WHO-5 score less than 50 ($TG1_{WHO-5}$) or LAR score less than 0.75 ($TG2_{LAR}$) on the average probability of whether a student self-refers to available student support services in the university; that is, to either CMH or SilverCloud self-help services, at the end of Semester 1 (December 2024). Students in the treatment groups who were not identified as at risk receive a generic email. Students in the control group receive no contact.

Regression analysis was also conducted to examine the impact of the interventions on the pre-specified secondary outcomes. As there are five secondary outcomes with comparisons between treatments as well as with the control, we conduct adjustments for multiple comparisons to the treatment estimates within the secondary outcomes, using the Benjamini-Hochberg step up procedure.

Finally, the direct impact of the intervention on student engagement was examined, specifically whether students clicked the links embedded in the intervention email. Since the control group were not emailed, the analysis was focused on which identification method ($TG1_{WHO-5}$ or $TG2_{LAR}$) was more effective in encouraging students to engage with the email.

In summary, our analysis tested three hypotheses for each outcome, specifically.

- **Hypothesis 1:** There will be a higher proportion of students allocated to treatment group 1, whose wellbeing is tracked via WHO-5, who self-refer to the support services on campus and who engage with email content (as captured by outcomes listed in Table 3) as compared to the students in the control condition
- **Hypothesis 2:** There will be a higher proportion of students allocated to treatment group 2, whose wellbeing is tracked via LAR, who self-refer to the support services on campus and who engage with email content (as captured by outcomes listed in Table 3) as compared to the students in the control condition

- **Hypothesis 3** There will be a higher proportion of students allocated treatment group 2, whose wellbeing is tracked via LAR to the support services on campus and who engage with email content (as captured by outcomes listed in Table 3) as compared to the students in treatment group 1, whose wellbeing is tracked via WHO-5.

Since the trial comprises of two-stage randomisation, a causal impact relies on two steps: (1) that the identifying mechanism successfully identifies and targets individuals who will benefit overall from the 'nudge' emails, and (2) that the emails, overall, are effective in encouraging engagement. The models overleaf estimate the combined impact of these two stages of potential causal effect.

Analysis was conducted in R using ordinary least squares (OLS) regression with robust standard errors with the following specification:

$$Y_i = \alpha + \beta_{1,2} D_i + \beta_{3:n} \boldsymbol{X_i} + \gamma_i + \epsilon_i$$

Where:

- Y_i is the outcome (see Table 3) of individual i;
- D_i is their randomised allocation (to $TG1_{WHO\text{-}5}$, $TG2_{LAR}$, or Control);
- $\boldsymbol{X_i}$ is a vector of individual-level covariates (see Section 4.3.2);
- ϵ_i is a Huber-White robust individual-level standard error.

For the continuation prediction score (secondary outcome), an additional term, Y_{0i} was included in the left-hand side of the equation to capture the baseline level of this score. Analysis of the secondary outcome measure, email click-through rates, also followed this analytical structure but without the inclusion of the control group and hence was only tested for the third hypothesis above.

This analysis has been conducted on an intention to treat basis on the sample of 6,001 participants sampled following the iterative process outlined in Section 4.2.4.

4.4.4 Multiple comparisons

The purpose of adjustment for multiple comparisons is to reduce the Type I (false positive) error rate arising due to the number of analyses that are carried out. In this study, per the protocol, adjustments for multiple comparisons were conducted using

the Benjamini–Hochberg False Discovery Rate method. In this method, p-values are first ordered from smallest to largest and then adjusted based on their rank, which controls the expected proportion of false positives across all statistical tests conducted.

4.4.5 Missing data and imputations

Missingness on pre-treatment covariates (including baseline outcomes) were assumed to be Missing Completely at Random. Following the pre-registered analysis protocol, since missingness on any given covariate was less than 5% complete case analysis was used. Further, attrition for outcomes collected via NU's management information was close to zero as all student records systems were linked using the same student ID. Thus, analysis was based on complete cases.

4.4.6 Robustness tests

The primary specification (see Section 4.4.3) was rerun including and excluding covariates that are available, including running a model excluding all covariates. As five of the outcomes of interest are binary, we also re-ran the primary analysis specification using a binary logistic (logit) estimator. We also examined how sensitive the treatment effect is to baseline learning analytics score. This is done by including an interaction between the treatment assignment and the baseline score of LAR. This specification allows us to explore whether the treatment impact differs across students with varying baseline performance. This provides a more nuanced and robust assessment of the intervention's effect by examining whether baseline scores are correlated with treatment responsiveness, which could otherwise bias the results.

We have reported these results in Appendix D.

4.4.7 Heterogeneity analysis

Subgroup analysis was conducted across i) gender, ii) first in family, and iii) faculty. These subgroups were selected for analysis because they represent key demographic and contextual factors that were expected to potentially influence the intervention's effectiveness. For instance, first in family is a good indicator of the cultural/social factors that may influence their engagement within an academic environment and, consequently, their response to the intervention. Gender differences in student engagement may be influenced by factors such as social expectations, access to resources, personal interests, and peer support. For example, female students may face additional academic and personal responsibilities, potentially lowering their engagement in university, making it valuable to examine whether the intervention had differential effects. Finally, analysing the impact by faculty is crucial for understanding whether academic culture play a role in their engagement and thus the association between the intervention and engagement levels.

For the subgroup analysis, we use the full sample of randomised participants (*n*=13,122). This has been investigated through the inclusion of relevant interaction terms in our models.

NU were also interested in exploring the differential effects on click-through rates based on whether students are first in family to attend university and based on their study status (new or continuing). The estimates have been presented in Appendix D.

4.4.8 Exploratory analysis

NU conducted a Wald test based on the full sample, including those scoped out of the RCT (that is, those participants who do not fall within the 6,001-participant analytical strategy) due to high scores on both WHO-5 and LAR. This was aimed to understand the effects of the interventions conditional on receiving it.

NU was interested in exploring the direct effect of receiving the nudge email on the primary outcome. To do this, we subsetted participants in $TG1_{WHO\text{-}5}$ and the control whose WHO-5 scores were below 50, and compared their outcomes, then did the same for participants in $TG2_{LAR}$ and the control whose LAR scores were below 75%.

NU also wanted to understand whether being assigned to a treatment group, regardless of which treatment received, has a differential impact on self-referral to support services compared to being assigned to the control group. To test this, we bundled together students in $TG1_{WHO\text{-}5}$ and $TG2_{LAR}$, classifying them as treated, and compared their self-referral to support services' status with that of students in the control group.

4.5 Deviations from study protocol

To the best of our knowledge, based on the data provided by NU, the intervention was delivered as intended. Overall, the analysis was conducted following the pre-registered trial protocol.

Nevertheless, two elements are worth highlighting when it comes to inclusion of covariates. First, the selection of covariates for inclusion in the model was intended to be dynamic, depending on what NU was able to provide without de-anonymising the data. Second, the protocol specified a campus fixed effect would be included in the analytical specification. However, including this variable in practice meant that many others were excluded as NU was not able to provide this field without risking de-anonymising the data by creating cells with too few cases in them. The decision was therefore taken not to provide and this variable.

The following covariates were ultimately included in the analytical specification, which are also included Table 4: faculty, ethnicity, first in family, gender, mature student status, and student status (new/continuing).

5. Results

5.1 Descriptive statistics

The full sample consists of 13,122 participants, with 4,367 in treatment group 1 ($TG1_{WHO-5}$), 4,367 in treatment group 2 ($TG2_{LAR}$) and 4,388 assigned to the control group.

There were 692 students from $TG1_{WHO-5}$ who fell below the WHO-5 score of 50, while 421 from $TG2_{LAR}$ fell below a LAR of 0.75 (75%). Students from each of these groups were further randomised to receive one of the two nudge emails, signposting to one-to-one counselling or online self-help. Those who did not fall below the threshold on either treatment-identification method received a generic signposting email. Those in the control group received no intervention.

Appendix C presents the key demographic variables—including gender, ethnicity, age group, study stage, faculty, and first-in-family status—for the full sample, broken down by each subgroup of the variables listed, along with the distribution within these subgroups by treatment allocation (see Appendix C: Descriptive analysis

Table 28). The analysis was conducted to summarise sample characteristics and set the baseline context for interpreting subsequent results.

5.2 Balance checks

To interpret a difference between the treated and control groups as a causal effect, we must assume that, in the absence of treatment, the groups would have been balanced in expectation on the outcome of interest. This is an inherently untestable assumption, as we cannot observe the untreated outcome for those who received the treatment. Instead, we assess balance on observable characteristics to evaluate whether randomisation was successful. Balance checks help determine whether the groups are statistically similar on average with respect to observed covariates. This is important because the covariates included in the regression model are used to adjust for any baseline differences. Balance checks therefore help assess whether those differences are small and of a nature such that regression adjustment can reasonably account for them.

Balance checks were conducted on the following observable characteristics: faculty, ethnicity, gender, age group, first in family to receive university education, and student study status (see Table 5). We considered that there is an imbalance on a covariate if there is more than a 5 percent difference in the proportions of respondents in each category for the binary variables. This criterion was not exceeded by any of the tested characteristics, which increases confidence that any findings are not due to baseline differences between the groups.

In general, as sample stratification was carried out using WHO-5 and LAR score (see Section 4.2.3) there was a risk that randomisation would not result in balanced treatment and control groups across variables commonly used in stratification such as gender, ethnicity, and department. However, as presented in Appendix C, the randomisation has resulted in treatment and control groups that are balanced across gender, ethnicity, study stage, faculty, and family history of higher education.

Table 5: Balance checks on key socio-demographic variables

Variable	Control	$TG1_{WHO-5}$	$TG2_{LAR}$	Absolute Difference 1	Absolute Difference 2
Faculty (Percentage in ADSS)	14.9	15.3	15.7	0.4	0.8
Faculty (Percentage in BL)	27.5	28.1	27.5	0.6	0.0
Faculty (Percentage in EE)	21.7	20.9	20.5	0.8	1.2
Faculty (Percentage in HLS)	35.9	35.7	36.3	0.2	0.4
Ethnicity (Percentage of White)	65.6	66.1	66.2	0.5	0.6
Gender (Percentage of Male)	43.9	45.0	44.3	1.1	0.8
Age group (Percentage of <=21 years)	60.7	61.2	59.5	0.5	1.2
First in family (Percentage of students who respond "Yes")	36.5	36.9	37.4	0.4	0.9
Student status (Percentage of Continuing)	52.6	52.9	53.1	0.3	0.5

Difference 1 refers to the difference in the proportions between control and $TG1_{WHO-5}$
Difference 2 refers to the difference in the proportions between control and $TG2_{LAR}$

5.3 Results of impact analysis

Table 6 below provides the outputs of this analysis for all outcomes, using the primary specification outlined in Section 4.4.3.

Table 6: Regression models, all outcomes

	Primary		Secondary								
	Self-referral to support services		LAR		Self-referral to counselling		Self-referral to SilverCloud		Withdrawal		Click-throughs
Model	*(1)*	*(2)*	*(1)*	*(2)*	*(1)*	*(2)*	*(1)*	*(2)*	*(1)*	*(2)*	*(2)*
(Intercept)	0.102 (0.076; 0.127) *	0.102 (0.072; 0.132) *	0.331 (0.295; 0.367) *	0.337 (0.294; 0.381) *	0.099 (0.074; 0.124) *	0.100 (0.070; 0.130) *	0.004 (-0.003; 0.012)	0.002 (-0.007; 0.011)	0.038 (0.021; 0.055) *	0.045 (0.024; 0.067) *	0.041 (0.020; 0.063) *
TG1	0.004 (-0.010; 0.019)		0.002 (-0.006; 0.011)		0.005 (-0.010; 0.019)		0.000 (-0.004; 0.005)		0.008 (-0.003; 0.019)		
TG2	-0.004 (-0.018; 0.011)	-0.008 (-0.022; 0.007)	0.002 (-0.007; 0.010)	-0.001 (-0.009; 0.008)	-0.003 (-0.017; 0.011)	-0.007 (-0.021; 0.007)	-0.001 (-0.005; 0.004)	-0.001 (-0.005; 0.004)	0.007 (-0.004; 0.018)	-0.001 (-0.012; 0.011)	-0.009 (-0.018; 0.001)
Faculty = BL	-0.043 (-0.065; -0.022) *	-0.039 (-0.066; -0.012) *	0.027 (0.015; 0.039) *	0.025 (0.010; 0.039) *	-0.042 (-0.063; -0.020) *	-0.038 (-0.064; -0.012) *	-0.006 (-0.012; 0.001)	-0.004 (-0.012; 0.004)	-0.006 (-0.022; 0.009)	0.002 (-0.017; 0.021)	-0.024 (-0.042; -0.005) *
Faculty = EE	-0.034 (-0.058; -0.011) *	-0.028 (-0.057; 0.001)	-0.004 (-0.017; 0.010)	-0.009 (-0.026; 0.007)	-0.035 (-0.058; -0.012) *	-0.030 (-0.059; -0.002) *	-0.003 (-0.010, 0.005)	0.000 (-0.009, 0.010)	-0.015 (-0.031, 0.001)	-0.012 (-0.032, 0.007)	-0.014 (-0.036, 0.007)
Faculty = HLS	-0.026 (-0.048, -0.005) *	-0.032 (-0.058, -0.007) *	0.007 (-0.005, 0.019)	0.002 (-0.012, 0.017)	-0.028 (-0.048, -0.007) *	-0.031 (-0.056, -0.006) *	-0.001 (-0.008, 0.006)	-0.003 (-0.011, 0.005)	0.000 (-0.014, 0.015)	0.007 (-0.011, 0.024)	-0.030 (-0.048, -0.012) *
Ethnicity = Non-white	-0.001 (-0.017, 0.015)	-0.004 (-0.025, 0.017)	0.015 (0.005, 0.025) *	0.017 (0.005, 0.029) *	0.001 (-0.015, 0.017)	-0.002 (-0.023, 0.019)	-0.001 (-0.006, 0.003)	0.000 (-0.006, 0.006)	-0.006 (-0.017, 0.006)	-0.006 (-0.021, 0.009)	0.001 (-0.012, 0.015)
Ethnicity = Unknown	-0.009 (-0.046, 0.028)	-0.005 (-0.052, 0.041)	0.016 (-0.017, 0.049)	0.006 (-0.031, 0.044)	-0.006 (-0.042, 0.031)	-0.002 (-0.048, 0.044)	-0.005 (-0.008, -0.002) *	-0.005 (-0.008, -0.001) *	-0.001 (-0.033, 0.032)	-0.001 (-0.042, 0.039)	-0.013 (-0.038, 0.012)
FIF = Unknown	-0.018 (-0.036, -0.001) *	-0.014 (-0.037, 0.010)	0.020 (0.008, 0.031) *	0.010 (-0.004, 0.023)	-0.020 (-0.037, -0.003) *	-0.015 (-0.038, 0.009)	-0.002 (-0.006, 0.003)	-0.003 (-0.007, 0.001)	-0.018 (-0.031, -0.005) *	-0.022 (-0.040, -0.005) *	-0.018 (-0.033, -0.003) *
FIF = Yes	0.011 (-0.005, 0.027)	0.010 (-0.010, 0.029)	0.002 (-0.007, 0.010)	-0.004 (-0.014, 0.007)	0.009 (-0.006, 0.025)	0.007 (-0.012, 0.026)	0.001 (-0.004, 0.007)	0.005 (-0.002, 0.011)	-0.006 (-0.018, 0.006)	-0.010 (-0.025, 0.005)	-0.005 (-0.018, 0.008)
Gender = Female	0.008 (-0.006, 0.021)	0.012 (-0.005, 0.028)	0.000 (-0.008, 0.008)	-0.003 (-0.014, 0.007)	0.005 (-0.008, 0.018)	0.007 (-0.009, 0.023)	0.005 (0.002, 0.009) *	0.007 (0.002, 0.012) *	-0.013 (-0.023, -0.003) *	-0.018 (-0.031, -0.004) *	0.014 (0.003, 0.025) *
Above 21 = Yes	-0.020 (-0.033, -0.008) *	-0.024 (-0.039, -0.008) *	-0.032 (-0.040, -0.024) *	-0.028 (-0.038, -0.019) *	-0.018 (-0.031, -0.006) *	-0.021 (-0.037, -0.006) *	0.000 (-0.004, 0.004)	0.000 (-0.004, 0.005)	0.011 (0.000, 0.021) *	0.009 (-0.004, 0.022)	0.015 (0.004, 0.026) *
Student status = New	-0.013 (-0.025, -0.001) *	-0.009 (-0.024, 0.006)	0.023 (0.016, 0.031) *	0.026 (0.018, 0.035) *	-0.013 (-0.025, -0.001) *	-0.008 (-0.023, 0.007)	0.002 (-0.002, 0.006)	0.001 (-0.003, 0.006)	0.015 (0.005, 0.024) *	0.018 (0.006, 0.031) *	0.003 (-0.008, 0.013)
R^2	0.015	0.013	0.420	0.432	0.014	0.011	0.004	0.005	0.006	0.007	0.010
Adj. R^2	0.013	0.011	0.424	0.430	0.012	0.009	0.002	0.002	0.004	0.004	0.007
Num. obs.	6001	4001	5845	3885	6001	4001	6001	4001	6001	4001	4001

*** p < 0.001; ** p < 0.01; * p < 0.05. 95% confidence intervals provided in brackets. LAR is coded to run from 0 – 1, representing likelihood of withdrawal of 0% to 100%. Model (1) compares each group to the control; Model (2) compares $TG2_{LAR}$ to $TG1_{WHO\text{-}5}$ (excluding control from the sample). Regression tables of all analytical specifications provided in Appendix D

5.3.1 Results of primary outcome analysis

There are no significant impacts of the method of identification of at-risk students on the average probability of students self-referring to support services in the university. The estimates are very small (0.4 percentage points for both the groups compared to control group) and none of the estimates are statistically significant.

5.3.2 Results of secondary outcome analysis

LAR (continuation prediction score)

There is no evidence of an impact of identification method on the LAR score. The coefficients are very small, representing a potential effect of 0.2 percentage points for both the groups compared to control group. No statistically significant effects are detected.

Self-referral to counselling services

There is no evidence of an impact of identification method on self-referral to counselling services, either SilverCloud or on-campus services (0.5 percentage points and -0.3 percentage points, for $TG1_{WHO-5}$ and $TG2_{LAR}$ respectively). No statistically significant effects are found.

Withdrawal

There is no evidence of an impact of identification method on the likelihood of withdrawal. As with other outcomes, the coefficients are extremely small (0.8 percentage points and 0.7 percentage points for $TG1_{WHO-5}$ and $TG2_{LAR}$ respectively). None of the effects were found to be statistically significant.

Email click-through rate

The coefficient from an OLS model for $TG2_{LAR}$ is not statistically significant.

A binary logistic regression model (Model 3 in Appendix D, Regression 7) revealed a directionally significant difference between $TG1_{WHO-5}$ and $TG2_{LAR}$, with the coefficient for $TG2_{LAR}$ estimated at 0.339, significant at the 10% level. This coefficient indicates that the odds of clicking an email were lower for students in $TG2_{LAR}$ compared to $TG1_{WHO-5}$. The sign and magnitude of the estimate suggest a disadvantage for $TG2_{LAR}$ in terms of click-through probability.

This difference could arise due to two possible mechanisms. First, the students identified via the WHO-5 score (TG1) may inherently be more responsive to emails. Because WHO-5 and LAR scores classify students differently, the subset of students falling below the <50 threshold and therefore eligible for personalised nudges may

differ in behavioural characteristics that relate to engagement. If so, this could imply that the WHO-5 score is a more effective tool compared to the LAR score when it comes to identifying students with poor wellbeing who are also more likely to engage with support emails.

Second, because the intervention uses a multi-stage design, the difference may partly reflect variation in the type of email sent across arms. In $TG1_{WHO-5}$, a larger number of students (693 out of 4,367) received one of the two personalised nudge emails, whereas in $TG2_{LAR}$ only 135 out of 4,367 received such personalised emails. If personalised emails have higher intrinsic click-through rates than generic emails, the higher overall click-through rates in $TG1_{WHO-5}$ may simply reflect this difference in composition rather than differences in behavioural responsiveness.

Therefore, the observed effect may reflect an interplay between (a) the identification method (WHO-5 vs LAR) and (b) *the* distribution of email types that each rule produces. A full decomposition of these pathways is beyond the scope of the present analysis, but the evidence suggests that the interaction of both factors may contribute to the TG1–TG2 difference.

5.3.3 Results of heterogeneity analysis

The heterogeneity analysis indicated that were no significant effects of the intervention on whether or not a student self refers to student support services (see Table 7).

Table 7: Heterogeneity Analyses

Subgroup	Treatment Group	Interaction Estimate	Standard Error	p value
Gender = Female (ref: Male)	1	-0.003	0.009	0.720
	2	0.004	0.008	0.665
First in Family = Yes (ref: No)	1	0.009	0.012	0.428
	2	0.012	0.011	0.299
First in Family = Unknown (ref: No)	1	0.001	0.010	0.955
	2	0.009	0.010	0.356
Faculty = BL (ref: ADSS)	1	0.004	0.016	0.807
	2	0.007	0.015	0.644
Faculty = EE (ref: ADSS)	1	0.004	0.017	0.829
	2	-0.002	0.016	0.884

Subgroup	Treatment Group	Interaction Estimate	Standard Error	p value
Faculty = HLS (ref: ADSS)	1	-0.011	0.017	0.489
	2	-0.006	0.016	0.718

*** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. Covariates include gender, age group, ethnicity, first in family indicator, faculty and student stage. Full regression tables provided in Appendix D.

5.3.4 Results of exploratory analysis

The direct effect of receiving the nudge email on the primary outcome was not found to be statistically significant (at the 5% level of significance) between those in the relevant treatment group and the control. To do this, we subsetted participants in $TG1_{WHO\text{-}5}$ and the control group whose WHO-5 scores were below 50 and compared their outcomes. Then we did the same for participants in $TG2_{LAR}$ and the control whose LAR scores were below 75%. In both cases, the rate of engagement on the primary outcome (self-referral to one-to-one Counselling or SilverCloud self-help) is not significantly different.

Further, NU also wanted to understand whether being assigned to a treatment group, regardless of which treatment received, has a differential impact on self-referral to support services compared to being assigned to the control group. To test this, we bundled together students in $TG1_{WHO\text{-}5}$ and $TG2_{LAR}$, classifying them as treated, and compared their self-referral to support services' status with that of students in the control group. No statistically significant results (at the 5% level of significance) were found between being assigned to treatment group, (regardless of which treatment received), compared to being assigned to the control group, with respect to self-referral to support services.

5.3.5 Review of study minimum detectable effect size

At the trial design stage, MDES calculations were conducted using the pwr package in R to determine the minimum effect size that the study would be powered to detect, based on the following assumptions:

- Significance level: 0.05
- Power: 0.8

Power calculations were conducted for a range of scenarios, focused on different levels of need (that is, different proportions of people triggering the need for support under the two conditions), and different sizes of sample drawn from within the 16,800 participants who were expected to make up the whole sample.

As discussed in Section 4.2.2, this trial faced a trade-off between sample size and the proportion of the sample that is treated. While increasing sample size generally improves statistical power, it also dilutes the treatment effect if the number of treated units is held constant — leading to attenuation that reduces the observed effect size. Because statistical power increases with diminishing returns as sample size grows, and attenuation of the treatment effect increases roughly linearly, a smaller sample (with higher treatment density) may be more effective to detect a given effect. This, in turn, must be traded off for the need for the intervention to be discriminant in both groups – that is, not everyone assigned to each group at random should be treated, otherwise a crucial part of the intervention, that is to test the effectiveness of identifying and appropriately targeting, itself is lost.

Accounting for the multiple trade-offs, between sample size and the proportion of the sample that is treated, we designed the trial in a way such that all the in-scope participants (n = 16,800) would be randomised to maximise sample size, but primary analysis would be conducted on a discriminant group of 6,000 participants. The process of creating the discriminant group has been described in Section 0.

Table 8 below compares the ex-ante and ex-post power calculations.

Table 8: MDES calculations for given sample size for pre- and-post trial stage

	Sample size	Size of $TG1_{WHO-5}$	Size of $TG2_{LAR}$	Size of control group	MDES (Cohen's d)	Effect type (unadjusted/attenuated)
Ex-ante	16,800	5,600	5,600	5,600	0.05	Unadjusted for number of treated individuals
					0.30	Attenuated effect when 1,000 participants treated per group
					0.19	Attenuated effect when 1,500 participants treated per group
	6,000	2,000	2,000	2,000	0.09	Unadjusted for number of treated individuals
					0.18	Attenuated effect when 1,000 participants treated per group
					0.11	Attenuated effect when 1,500 participants treated per group
Ex-post	13,122[1]	4,367	4,367	4,388	0.06	Unadjusted for number of treated individuals
					0.38	Attenuated effect when 692 treated participants in TG1
					0.62	Attenuated effect when 421 treated participants in TG2
	6,001[2]	2,010	1,991	2,000	0.09	Unadjusted for number of treated individuals
					0.26	Attenuated effect when 692 treated participants in TG1
					0.42	Attenuated effect when 421 treated participants in TG2

[1]MDES computed assuming size per arm = 4,374

[2]MDES computed assuming size per arm = 2,000

Following trial implementation, the actual sample size achieved was 13,122 while the expected sample size was 16,800. Nonetheless, the effect sizes observed for most outcomes in the trial are estimated using the discriminant group, the sample size of which was 6,001. This was in line with the expected sample sizes to be achieved. However, as shown in Table 8, fewer students than expected triggered a high-risk status in the trial, and thus received the nudge emails, lowering the power of the trial to detect smaller effect sizes. The total treated equalled 692 in $TG1_{WHO-5}$ and 421 in $TG2_{LAR}$ while the expectation was between 1,000-1,500 students. While the initial design allowed detection of small effects (Cohen's d = 0.1–0.2), the smaller than expected treatment uptake increased the MDES to between 0.3–0.4 - corresponding to moderate effects. It must be noted that, since the intervention targets behavioural outcomes and is evaluated over a short timeframe, it is likely to produce only very small effects. This limits the design's ability to detect meaningful treatment effects, if they exist.

5.4 Summary of impact evaluation findings

5.4.1 Summary of caveats and limitations: impact evaluation

Table 9 presents a summary of the risk of bias assessment conducted using the Risk of Bias 2 (RoB 2) tool, which is designed to evaluate individually randomized control trials. The tool assesses five domains of potential bias that may affect the validity of the trial's results. Each domain is judged as "Low risk", "Some concerns", or "High risk" of bias, and an overall risk of bias judgment is also provided for each study. This summary helps to transparently report the methodological quality and credibility of the included trials.

Table 9: Summary of assessment of risk of bias assessment using RoB 2 tool for individually randomised parallel group trial

Domain	Comment	Risk of bias
Randomisation sequence and efficacy	The study employed a robust randomisation process with stratification, ensuring that the assignment of participants to treatment and control groups was well balanced and free from selection bias.	Low
Deviations from intended intervention	There were no known deviations from the intended intervention. Monitoring data tracked students' engagement rates and have been referenced in the IPE section. Overall, it was found that email delivery was successful. The implementation followed the planned delivery protocol. No departures from the scheduled content, timing, or mode of delivery were reported by delivery staff, indicating a low likelihood of major deviations.	Low

Domain	Comment	Risk of bias
Missing outcome data	Most of the outcomes' data was sourced from NU's management information resulting in near-zero attrition; therefore, for the primary and secondary outcomes, risk of bias arising from missing data is low.	Low
Measurement of outcome data	Outcome data were drawn from NU's management information, with consistent procedures applied across all participants. There was no evidence of measurement bias limiting possibility of systematic errors.	Low
Overall risk of bias score	In our view, the study was conducted as per the agreed protocol with randomisation conducted successfully and intervention delivered as expected. There exists no measurement bias or bias from missingness due to the nature of the data source, lending credibility and reliability to the results obtained from the trial.	Low

6. Implementation and process evaluation (IPE)

The aim of this IPE was to examine how two different methods of identifying students' wellbeing – (i) a self-reported survey and (ii) predictive learning analytics – influence students' engagement with email nudges and their propensity to subsequently seek help. The IPE also considered institutional factors which directly influence the ability to nudge in this way such as capturing students' consent to be identified and nudged, as well as staff perspectives on supporting the activity. Finally, quantitative data covering students' demographics, their engagement with the intervention and their academic engagement are presented.

6.1 Research questions

This IPE has been designed to address the following four research questions:

- RQ1: What is the impact of the consent process and how does it influence the different methods of identifying students with poor wellbeing who require a proactive intervention?
- RQ2: How do students engage with wellbeing nudges?
- RQ3: Does proactively offering wellbeing support have an impact on students' propensity to seek help?
- RQ4: What are staff views on the different approaches to identifying students in need of support?

Each of the research questions are connected to specific IPE dimensions (see Table 10) from the RE-AIM framework (Glasgow et al., 1999) which has five domains; implementation, reach, adoption, effectiveness and maintenance. The dimensions of implementation, reach, and maintenance map onto TASO's IPE dimensions of adherence, reach, and sustainability.

Table 10: IPE dimension and relevance

IPE dimension	Relevance	Research Question (RQ)
Reach	The extent the nudge intervention and associated wellbeing support was accessed by or accessible to the intended population and what characteristics were distinctive of those that did or did not engage.	RQ1, RQ2
Effectiveness	The positive and negative outcomes of the nudge intervention.	RQ2, RQ3

IPE dimension	Relevance	Research Question (RQ)
Adoption	The proportion of settings and providers of the wellbeing support who are willing to initiate a nudge intervention to support services and why.	RQ4
Implementation (Adherence)	The extent the intervention was embedded and delivered as intended.	RQ4
Maintenance (Sustainability)	The extent the nudge intervention was (or could be) embedded and sustained as part of routine practice. As this intervention was a single delivery, actual maintenance cannot be discussed. However, the potential for embedding the intervention is interpreted from the data.	RQ3

6.2 IPE design and framework

The IPE is a mixed-methods approach to answer the research questions and provide understanding of the mechanisms which underpin how and why the intervention is (or is not) effective.

The IPE is informed by a combination of qualitative and quantitative information about the intervention. These include focus groups with staff involved in delivering the intervention, student data from the management information systems (student records system, the CRM system that handles email communication, and the WHO-5 survey), and student risk data from the learning analytics system (the predicted likelihood of a student continuing or completing their education). Together these sources provide a comprehensive view of how the intervention was delivered, how students interacted with the nudges, the sustainability of the intervention and any contextual factors that may improve its effectiveness

Table 11: IPE framework

IPE dimension	Audience	Source of data	Data analysis method	RQ
Reach	Students	Consent rate from SITs system	Quantitative-Descriptive Statistics	1, 2
		Open/click-through rate of each URL in the nudge email from the CRM		
	Students	Distribution of risk data for risk cut off from the learning analytics system		

IPE dimension	Audience	Source of data	Data analysis method	RQ
	Students	Distribution of risk data for risk cut off from the WHO-5 survey (stored in SITS)	Quantitative-Descriptive Statistics	
Effectiveness	Students	Reply rate from the Outlook email server		2, 3
	Students	Change of Circumstances (CHoC) Data to ascertain whether students seek to postpone/break or leave their study from student record system.		
	Students	CORE-10 data from the CRM to ascertain whether the students who engage with services post-intervention are clinically in need of the support they sought		
Adoption and Maintenance	Staff	Semi-structured focus group	Qualitative	3, 4
		7-point NOAMD-Lite Survey (see Appendix E: Data collection tools for IPE)	Quantitative-Descriptive Statistics	
Implementation	Students	Delivery rate from management information system (CRM)	Quantitative analysis – (binary) logistic regression	4

The aim of the nudge is to encourage students to come forward who need help but may not typically ask for it. We used an implementation theory (Normalization process theory; NPT) to explore the implementation of the project from the perspective of staff who are directly involved in supporting students once they come forward and seek help. The aim is to understand whether the practice of data-informed nudging works and to make recommendations for the embedding of this type of support in services.

NPT identifies four determinants of embedding (i.e. normalizing) complex interventions in practice (May & Finch, 2009) and is therefore well placed in understanding the viability of this intervention in university services:

1. **Coherence or sense making** - The extent stakeholders can make sense of a new intervention and differentiate the ways it deviates from established ways of working
2. **Cognitive participation or engagement –** The relational work that stakeholders do to build and sustain a community of practice around a new intervention/way of working.
3. **Collective action** - Operational work that people do to enact a set of practices that enable the intervention to happen.
4. **Reflexive monitoring** - Formal and informal appraisal of the benefits and costs of the intervention.

We also used the RE-AIM evaluation framework (Glasgow et al., 1999) to evaluate the intervention's success or failure in terms of the IPE dimensions in Table 10. In contrast to NPT, Glasgow's RE-AIM Framework (*ibid*) proposes five domains that can influence the implementation of new services across a range of stakeholders:

1. **Reach –** The absolute number, proportion, and representativeness of individuals who are willing to participate in an intervention.
2. **Effectiveness –** The impact of an intervention on important outcomes, including potential negative effects, quality of life, and economic outcomes.
3. **Adoption** - The absolute number, proportion, and representativeness of settings and intervention agents (people who deliver the programme) who are willing to initiate a programme.
4. **Implementation** - the extent to which the programme was delivered as intended and consistently, and its cost.
5. **Maintenance** - The extent to which an intervention becomes part of the routine organizational practices and policies. At the individual level, maintenance has been defined as the long-term effects of a programme on outcomes after six or more months after the most recent intervention contact. As this intervention was a single delivery, themes of maintenance were not discussed.

The RE-AIM Framework has been applied to understand intervention impact across a variety of healthcare settings and acknowledges the value of qualitative data to complement quantitative measures.

6.3 Data

As highlighted above, data for the IPE were collected through a range of sources of information for the five dimensions in Table 11:

- **A Semi-structured focus group** was conducted with staff to inform our response to RQ4 predominantly but also to contextualise findings from RQ1. The aim was to interview six staff involved in the implementation of the programme (see further information below).
- **A NOMAD-Lite survey** was conducted with staff. The Normalisation Measurement Development (NoMAD) tool (Finch et al., 2013) is a 23-item questionnaire designed for use in implementing complex healthcare interventions from the perspective of professionals working in the environment. Modified and scaled versions of the NoMAD have been found to be effective methods in health research (Satherley et al., 2025) and therefore 7 statements were taken and used not only as a guide for the focus group but also a simple data collection survey tool to prompt participants to think about the intervention from the perspective of their role.

Various data from different sources of **management information systems** were compiled for this analysis:

- **SITS:** Data relating to opt-in rates, CHoC, WHO-5 and enrolment status were taken from the student record system.
- **University marketing CRM:** students' engagement with email nudges (delivery, open, and click-through rates).
- **Student support CRM**: sign-ups to CMH support, along with responses to the CORE-10 on registration were downloaded from the student support CRM.
- **Analytics system**: NU uses a learning analytics dashboard (Illume, created by Civitas Learning) that draws data from a range of institutional systems to track how engaged a student is with their learning. These data include virtual learning environment (VLE) activity, attendance monitoring, submissions, grade performance, and building access. The dashboard generates a LAR score for the predicted likelihood of a student having a positive retention outcome (repeat, progress or complete) the following academic year (where 0% is very unlikely to have a positive outcome and 100% being very likely).

6.3.1 Sampling strategy

Staff Recruitment

Qualitative data were collected via an in-person focus group in November 2024, three weeks after an email nudge had been sent to students. The email wording related to the

focus group was provided by the research team in accordance with our ethics application; this included a short overview of the researchers, the project, the time commitment and venue. The invitation was made in collaboration with the manager of the CMH team who included this in her weekly team communication. Participants were invited purposively for their role in therapeutic delivery of mental health and wellbeing services within the counselling team of a UK university. Of 14 practitioners in the team, six attended the focus-group; attendance was based on the availability of staff members based on their duty rota and their willingness to participate.

Quantitative data

The total population included students who enrolled in the academic year 2024/25. The intervention population was students who consented to both their educational analytics data and their WHO-5 data being used to identify them for wellbeing support. Of the 24,575 students who were enrolled at the beginning of term, 86.8% (21,331) consented to the use of their educational analytics data and 61.5% (15,112) consented to the use of their WHO-5 data. In total, 13,122 (54.1%) students consented to being part of the RCT.

Table 12: Population levels of the different groups considered in the IPE

Sample population	Intended sample size	Achieved sample size
Staff interviews	6	6
Enrolled Students	28,000	24,575
EA sample: Consented to Educational Analytics (EA)	n/a	21,331 (86.8%)
WHO-5 sample: Consented to EA and WHO-5 data driven support	n/a	15,112 (61.5%)
Sample population: Consented to be part of the randomised controlled trial	16,800	13,122 (54.1%)

In addition to the top-level randomisation of students into $TG1_{WHO-5}$, $TG2_{LAR}$ and Control, students in $TG1_{WHO-5}$ and $TG2_{LAR}$ were further randomised into three groups which determined the content of the email they were sent. Randomisation was stratified by a median split of the two scores (WHO-5 and LAR), creating four strata. The code for this is presented in Appendix G(ii): Randomisation code for second level randomisation . Table 13 shows the breakdown of the trial population based on the email nudge they received.

Table 13: Population levels of the different emails sent

Sample population	Count		
	$TG1_{WHO-5}$	$TG2_{LAR}$	Control
No Nudge	0	0	4,388
Email 1 (*n* = 558) Subject: Signposting you to Wellbeing Support via one-to-one Support	347	211	0
Email 2 (*n* = 555) Subject: Signposting you to Wellbeing Support via SilverCloud	345	210	0
Email 3 (*n* = 7,621) Subject: Signposting you to Wellbeing Support on your Student Portal	3,675	3,946	0

6.3.2 Research material design

The materials used for the semi-structured focus group include a discussion guide (Appendix E) and participant information sheet (Appendix Eii).

- **Participant information sheet**: The participant information sheet provided clear and accessible details about the project, why staff have been contacted, and what participation in the study means for them and how they can opt out of the study if needed.
- **Focus group guide and survey:** The discussion guide/ survey tool was designed by selecting seven relevant items from the NOMAD survey (Appendix E).
- **Enrolment screen:** This screen was presented at enrolment and used to provide information and obtain consent for participation in the trial (Appendix Eiii).

6.3.3 Focus group conduct

The focus group was held in-person. It was semi-structured and guided by the seven-item guide to frame the discussion. Participants were allowed to speak freely and other questions were generated during discussion (Appendix E). It lasted approximately one hour and was held in a secure meeting room in the library on campus. The focus group then explored participants' answers to understand in more detail how they had responded. Finally, participants were asked to revisit their original responses at the end of the focus group so that any changes could be recorded. Data were captured using audio recordings and transcription software was used to transfer audio to word format for analysis.

6.3.4 Approach to maximising response rates

No incentives were used in the collection of data in this project. An email reminder was sent to the manager of the CMH team the day before the focus group as a reminder that staff would be joining the session to avoid any last-minute scheduling clashes.

6.4 Analytical strategy

The IPE employs a mixed-methods approach and uses both qualitative and quantitative data to answer the research questions.

Qualitative data from the focus group was analysed using a thematic template analysis (Brooks et al., 2015). Top-down coding followed the four determinants of NPT and the five dimensions of the RE-AIM framework (except Maintenance). Second-level coding identified exploratory themes within the NPT and RE-AIM components. Both the principal investigator and academic lead were involved in the analysis and were able to quality check each other's work. his was achieved by each carrying out separate transcriptions on a section of the recording and calibrating the data to ensure interpretation was standard.

A standard array of descriptive statistics was utilised to understand the quantitative data. Chi-squared tests were used to compare groups in the treated population in terms of consent percentages, click-through rate percentages, and so on.

6.5 Deviations from study protocol

The initial study protocol outlined that an experimental cluster analysis would be conducted to a) ascertain the optimal number of clusters from the whole population and b) assess the key characteristics of the clusters. Given the time constraints on the project, the cluster analysis was not progressed.

6.6 Results

6.6.1 RQ1: What are the impacts of the consent process and the different methods of identifying poor wellbeing on the students who experience the intervention?

The sample population – those eligible to be randomised as part of the trial – is dependent on which students consent to their analytics and wellbeing data being used to provide nudges, and the subset of those students who go on to consent to be part of the trial; see Figure 4. After the consent process there was a final step to remove 51 students who had withdrawn, taken a break from their studies or transferred course before the nudge in week four, and a further 126 students who subsequently withdrew

their consent. This meant the final trial population was 54.1% (n = 13,122) of the total population.

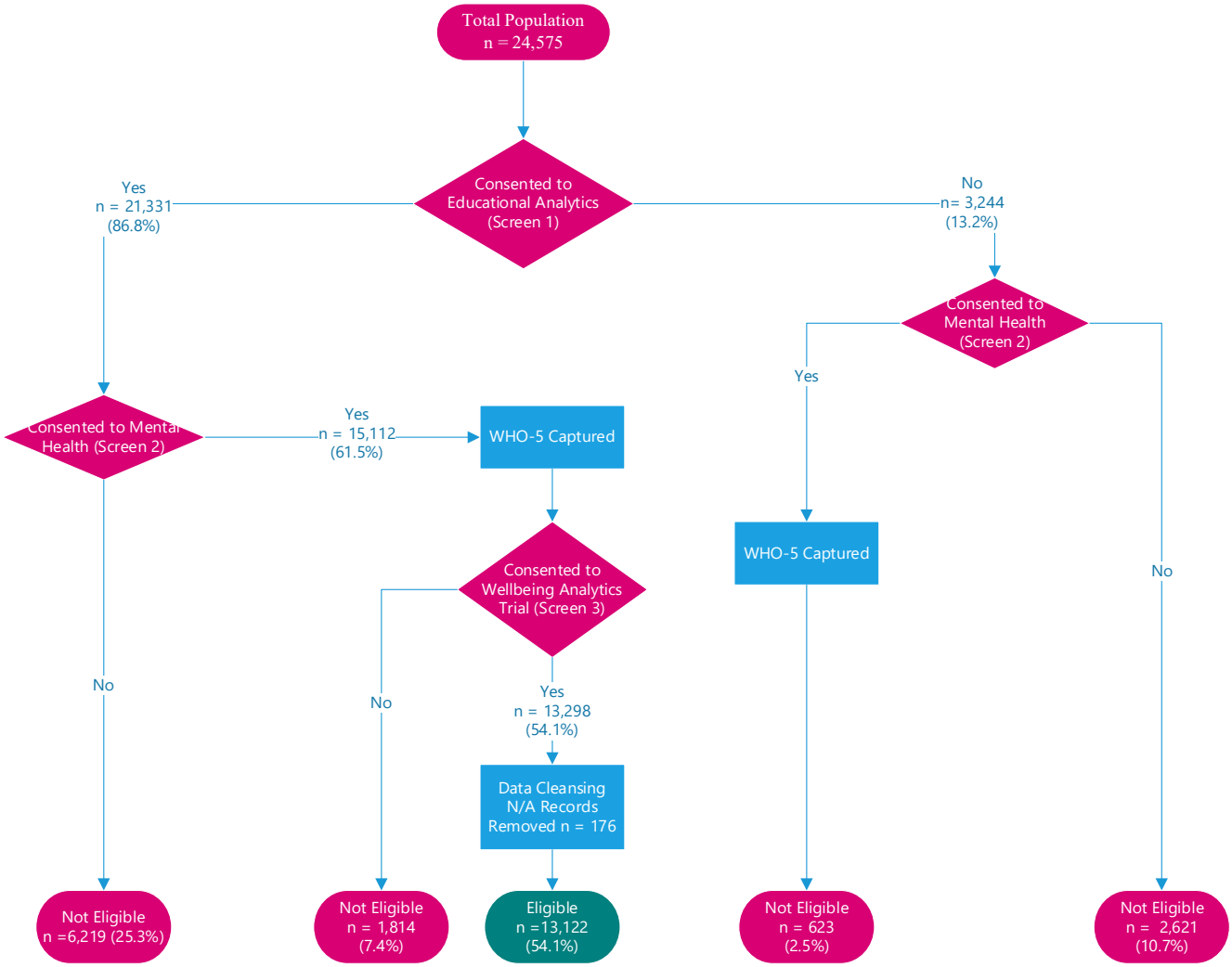


Figure 4: Flowchart showing consent-based approach

Relationship between gender and rates of consent

The relative levels of consent for each aspect of the trial split by gender is shown in Table 14. There is a significant relationship between gender and the likelihood of consenting to be part of the analytics, WHO-5 process, and the trial. Female students are more likely than males to sign up for EA support (88.9% vs. 84.4%, $\chi 2(1) = 109.72$, $p < 0.001$), wellbeing support (64.5% vs 58.1%, $\chi 2(1) = 103.19$, $p < 0.001$) and the present wellbeing trial (56.0% vs 50.4%, $\chi 2(1) = 78.31$, $p < 0.001$).

Table 14: The effect of the consent process on student demographics from all enrolled students to the trial sample (gender).

	Gender Count (%)			
Sample population	Male	Female	Other / Unknown	Total
Enrolled Students	11,564	12,936	75	24,575
EA sample: Consented to Educational Analytics (EA)	9,760 (84.4%)	11,506 (88.9%)	65 (86.7%)	21,331 (86.8%)
EA-WHO-5 sample: Consented to both EA and WHO-5	6,722 (58.1%)	8,339 (64.5%)	51 (68.0%)	15,112 (61.5%)
Sample population: Consented to be part of the research trial	5,825 (50.4%)	7,248 (56.0%)	49 (65.3%)	13,122 (54.1%)

Relationship between ethnicity and rates of consent

The relative levels of consent for each aspect of the trial split by ethnicity is shown in Table 15. There is also a significant relationship between ethnicity and the likelihood of consenting to be part of the analytics, WHO-5 process, and the trial. White students are significantly more likely than students from Black and Minority Ethnic (BAME) backgrounds to sign up for EA support (87.8% vs. 86.0%, $\chi2(1) = 13.187$, $p < 0.001$) but BAME students were more likely to consent than white student to the use of both analytics and the use of the WHO-5 (65.6% vs 60.2%, $\chi2(1) = 63.246$, $p < 0.001$). Students with unknown ethnicity status also had significantly lower consent rates for the trial (42.9%) than white students (51.4%; $\chi2(1) = 17.02$, $p < 0.001$) or BAME students (58.9%; $\chi2(1) = 115.37$, $p < 0.001$); these students may come through non-standard enrolment pathways such as short courses where ethnicity data capture has not been completed.

Table 15: The effect of the consent process on student demographics from all enrolled students to the trial sample (ethnicity).

	Ethnicity Count (%)			
Sample population	White	BAME	Other / Unknown	Total
Enrolled Students	16,648	7,300	627	24,575
EA sample: Consented to Educational Analytics (EA)	14609 (87.8%)	6281 (86.0%)	441 (70.3%)	21,331 (86.8%)
EA-WHO-5 sample: Consented to both EA and WHO-5	10017 (60.2%)	4789 (65.6%)	306 (48.8%)	15,112 (61.5%)
Sample population: Consented to be part of the research trial	8553 (51.4%)	4300 (58.9%)	269 (42.9%)	13,122 (54.1%)

Students who met the risk threshold

The number of students who met the risk threshold for the whole sample and by trial group (control or treatment) is in Table 16. The sample population is broken down by those who met only a single risk threshold (WHO-5 only or LAR only), or who met the risk threshold on both measures. Across all 13,122 students in the trial, 3,093 (24%) met the definition of at-risk due to their wellbeing survey score or analytics continuation prediction score. Of those, only 261 (8%) met the definition of at-risk by *both* measures.

Across the trial population 2,086 (15.9%) students met the threshold of poor wellbeing (WHO-5 less than 50), including the 261 students who met the risk threshold for both LAR and WHO-5. Of those students, 692 were assigned to $TG1_{WHO-5}$, 679 to $TG2_{LAR}$ and 715 to the control group (Table 29). The LAR score for identifying students at-risk was set at 75%; 1268 (9.7%) students in the trial population met that criterion of which 436 were assigned to $TG1_{WHO-5}$, 421 to $TG2_{LAR,}$ and 411 to the control group (Table 30).

Table 16: Treatment groups after randomization

	Control Group	TG1 (WHO-5)	TG2 (LAR)	Total Population
Population	4,388	4,367	4,367	13,122
At-risk – WHO-5 only	623 (14.2%)	598 (13.7%)	604 (13.8%)	1,825 (13.9%)
At-risk – LAR only	319 (7.3%)	342 (7.8%)	346 (7.9%)	1,007 (7.7%)
At-risk – LAR and WHO-5	92 2.1%)	94 (2.2%)	75 (1.7%)	261 (2.0%)
Not at-risk	3354 (76.4%)	3333 (76.3%)	3342 (76.5%)	10,029 (76.4%)
Average (Min, Max) of WHO-5	68.8 (0, 100)	69.2 (0, 100)	69.4, (0, 100)	69.1 (0, 100)
Average (Min, Max) of LAR (Pre Intervention)	88.8% (1.2%, 100.0%)	88.6% (1.7%, 100%)	88.7% (0.8%, 100%)	88.7% (0.8%, 100%)

Gender and risk threshold

The distribution of gender for the at-risk groups is in Table 17. Comparing the prevalence of any risk between male and female we see a significantly lower prevalence amongst the male population (20.5% vs 25.8%, $\chi^2(1) = 48.53$, $p < 0.001$). However, while females (17.0%) are significantly more likely than males (9.7%) to meet the definition of poor wellbeing according to WHO-5 ($\chi^2(1) = 143.31$, $p < 0.001$) they are significantly less likely to meet the LAR threshold (6.7% vs 8.9%; $\chi^2(1) = 21.0$, $p < 0.001$). There is no significant difference between males and females for meeting both risk criteria (1.9% vs 2.0%; $\chi^2(1) = 0.10$, $p = 0.7486$).

Despite the very small sample size students who identify as non-binary or who have refused to disclose their gender have a significantly higher rate of risk in the trial population (55.1%, $n = 27/49$) compared to students who have recorded a binary gender (23.5%, $n = 3{,}066 / 13{,}073$; $\chi^2(1) = 25.42$, $p < 0.001$) largely due to the higher rates (44.9%) of wellbeing risk (WHO-5) rather than learning analytics risk (4.1%).

Table 17: Population distribution of gender for different measures of risk

	Male	Female	Other / Unknown
Population	5,778	7,295	49
Any risk	1,187 (20.5%)	1,879 (25.8%)	27 (55.1%)
At-risk – WHO-5 only	562 (9.7%)	1,241 (17.0%)	22 (44.9%)
At-risk – LAR only	514 (8.9%)	491 (6.7%)	2 (4.1%)
At-risk – LAR and WHO-5	111 (1.9%)	147 (2.0%)	3 (6.1%)

Ethnicity and risk threshold

The distribution of ethnicity across at-risk groups is in Table 18.

Overall, white students are more likely than students who are not white (the combination of BAME and other ethnicities) to meet either one of the risk thresholds (24.8% vs 21.2%; $\chi^2(1)=21.31$, $p<0.001$). This is driven by white students being more likely to meet the WHO-5 threshold for poor wellbeing (15.9%) than for LAR risk (6.8%), $\chi^2(1) = 357.24$, $p <0.001$; this significant difference is not observed for students who are not white and who are therefore equally likely to meet one of the risk categories (10.0% for WHO-5 vs 9.4% for LAR; $\chi^2(1) = 0.8562$, $p = 0.352$).

Table 18: Population distribution of ethnicity for different measures of risk

	White	BAME	Other / Unknown	*Not White*
Population	8,651	4,212	259	4,471
Any risk	2146 (24.8%)	880 (20.9%)	67 (25.9%)	947 (21.2%)
At-risk – WHO-5 only	1,378 (15.9%)	420 (10.0%)	27 (10.4%)	446 (10.0%)
At-risk – LAR only	588 (6.8%)	386 (9.2%)	33 (12.7%)	419 (9.4%)
At-risk – LAR and WHO-5	180 (2.1%)	74 (1.8%)	7 (2.7%)	81 (1.8%)

Signposting to support and the different levels of risk

An additional level of randomisation was added in terms of the emails that were sent to students in the two treatment groups (see Table 19). The average WHO-5 scores and LAR scores were lower in the two groups who were nudged towards specific services (Counselling or SilverCloud) despite the range of scores being roughly equal in terms of the distribution of risk. Only 169 students were at risk for both WHO-5 and LAR.

Table 19: Emails sent to students in the treatment groups split by risk category.

	Email 1 Signposting to one-to-one Support	Email 2 Signposting to SilverCloud	Email 3: Signposting to Student Portal	Total Nudged Population
Population	558	555	7,621	8,734
At risk – WHO-5 only	304 (54.5%)	294 (53%)	604 (7.9%)	1202 (13.8%)
At-risk – LAR only	177 (31.7%)	169 (30.5%)	342 (4.5%)	688 (7.9%)
At-risk – LAR and WHO-5	77 (13.8%)	92 (16.6%)	0 (0%)	169 (1.9%)
Not at-risk	-	-	6675 (87.6%)	6675 (76.4%)
Average (Min, Max) of WHO-5	48.2 (0, 100)	48.0 (0, 100)	72.4 (0, 100)	69.3 (0, 100)
Average (Min, Max) of LAR (pre-intervention)	74.7% (0.8%, 100%)	74.0% (1.0%, 100%)	90.7% (1.7%, 100%)	88.6% (0.8%, 100%)

Students who register for counselling services are asked to complete the CORE-10 questionnaire. In total, 501 students completed the CORE-10; 338 had been nudged as part of the trial (338/ 8,734; 3.9%) while 163 students who were in the control group had not (163/ 4,388; 3.7%). The mean CORE-10 score for those students who had been nudged to one-to-one support (20.6, s.d. = 5.3) is lower than that for those in the control group (21.3, s.d. = 5.9). The theory of change stipulated that the nudge may bring forward students in need of support who may not have otherwise asked for it.

Firstly, we tested the relative likelihood of registering for one-to-one counselling services between students who received the different nudges. Table 20 shows the number of students who signed up to Counselling depending on the communication that was sent to them.

Table 20: Count of Counselling Registrations by email sent

	Email 1 Signposting to one-to-one support	Email 2 Signposting to SilverCloud
Total	558	556
Registered for Counselling	50 (8.96%)	33 (5.9%)

Although students directly signposted to one-to-one wellbeing support were more likely to register for counselling than those signposted to SilverCloud, the difference was not significant (8.96% vs 5.9%, $\chi 2(1) = 3.27$, $p = 0.07$).

Secondly, a comparison of means was conducted to ascertain whether those who completed the CORE-10 when they signed up for one-to-one support and who had been nudged had a lower risk score than those who were not nudged. The tests for normality (Shapiro-Wilk) and equal variance (Bartlett's test) confirmed the appropriateness of a t-test on the CORE-10 score. There is no significant difference between the mean CORE-10 score of the nudged students ($\mu = 20.6$) compared to the control group ($\mu = 21.3$); $t = 1.173$, $df = 499$, $p = 0.121$.

Staff perspectives

Staff participants in the focus groups recognised that the intervention was not engaging with the same cohort of students they typically would engage with and instead saw the intervention as targeting students at a lower threshold of symptoms and/or before their symptoms had escalated to a point which more required intensive support.

> *"We very rarely see students registering who are at the early stages. 'I've just noticed I've started to get a little bit more anxious and I don't want it to get….' We don't get that as much as we get people in a worst state. (P2)*

> *"I mean, yeah, they don't preempt." (P4)*

When asked whether they felt these students should be coming to the service, participants were in universal agreement that the nudging intervention was picking up students that warrant the service.

> *"'Cause it's a help seeking service no matter how we get to it. Whatever roots nudge or non nudged. It's legitimate" (P3)*

"It supports referrals coming in" (P7)

"It'd be definitely be different if were loads of referrals that actually weren't anything to do with mental health. Actually, they didn't need this, but the majority of them do." (P6)

6.6.2 RQ2: How do students engage with wellbeing nudges?

An initial quality check determined that 100% of the nudge emails were successfully delivered to the student. Overall, 17.7% (1544) of nudges were opened by students but there are differences in the open-, click-through- and reply-rates between treatment groups and which email was sent.

Email open rates

The data in Table 21 and in Figure 5 details the email open rates for each treatment group. Students in $TG1_{WHO-5}$ had a slightly higher open rate than those in $TG2_{LAR}$ (18.5% vs 16.8% respectively; $\chi^2(1) = 4.43$, $p = 0.04$). Compared with those in $TG2_{LAR}$, students in $TG1_{WHO-5}$ were more likely to open emails signposting one-to-one support (58.5% compared with 38.4%; $\chi^2(1) = 20.44$, $p < 0.001$). Although the rate of opening was higher for one-to-one support compared with SilverCloud (50.9% vs 46.8%) this was not found to be significant ($\chi^2(1) = 1.67$, $p = 0.197$). Generic emails signposting the standard wellbeing support had the lowest open rates overall (13.1%).

Table 21: Email open rate

	Email open rates					
	All Treatment Groups		$TG1_{WHO-5}$		$TG2_{LAR}$	
Any nudge (*n* = 8,734; TG1 = 4,367; TG2 = 4,367)	1544	17.7%	810	18.5%	734	16.8%
Email 1 (*n* = 558; TG1 = 347; TG2 = 211) Signposting to one-to-one Support	284	50.9%	203	58.5%	81	38.4%
Email 2 (*n* = 555; TG1 = 345; TG2 = 210) Signposting to SilverCloud	260	46.8%	170	49.3%	90	42.9%
Email 3 (*n* = 7,621; TG1 = 3,675; TG2 = 3,946) Signposting to Student Portal	1000	13.1%	437	11.9%	563	14.3%

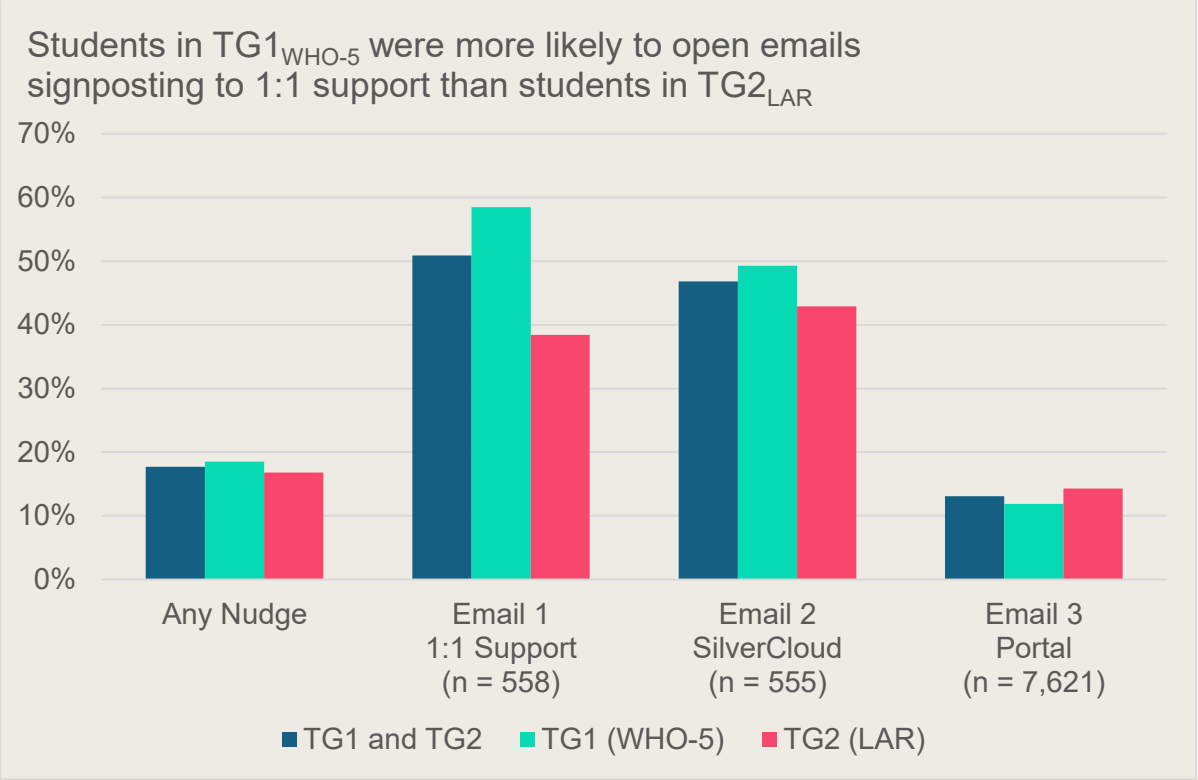


Figure 5: Email open rates for the three different kinds of email and for each treatment group.

Click-through and reply rates

The nudges sent in this study provided details about wellbeing support and direct links to one of a number of options. The click-through and reply rates for each email indicate the effectiveness of the nudges in signposting services, regardless of whether students went on to engage with those services. Only six students replied to the email and all replies were to acknowledge the email or thank the sender, so these are not considered further.

Click-through rates for each treatment group and email can be found in Table 22 and Figure 6. The overall click-through rate of the nudges was 2.1%. As with open rates, students in $TG1_{WHO-5}$ had a higher click-through rate than those in TG_{LAR} (2.4% vs 1.7%; $\chi^2(1)=4.72$, $p < 0.03$). Emails signposting one-to-one support had higher click-through rates than either SilverCloud (6.3% vs 3.4%; $\chi^2(1)=4.30$, $p = 0.04$) or generic wellbeing support (6.3% vs 1.7%; $\chi^2(1)=53.83$, $p < 0.001$).

Table 22: Click-through Rate

	Click-through rates					
	All Treatment Groups		$TG1_{WHO-5}$		$TG2_{LAR}$	
Any nudge (*n* = 8,734; TG1 = 4,367; TG2 = 4,367)	182	2.1%	106	2.4%	76	1.7%
Email 1 (*n* = 558; TG1 = 347; TG2 = 211) Signposting to one-to-one Support	35	6.3%	31	8.9%	4	1.9%
Email 2 (*n* = 555; TG1 = 345; TG2 = 210) Signposting to SilverCloud	19	3.4%	16	4.6%	3	1.4%
Email 3 (*n* = 7,621; TG1 = 3,675; TG2 = 3,946) Signposting to Student Portal	128	1.7%	59	1.6%	69	1.7%

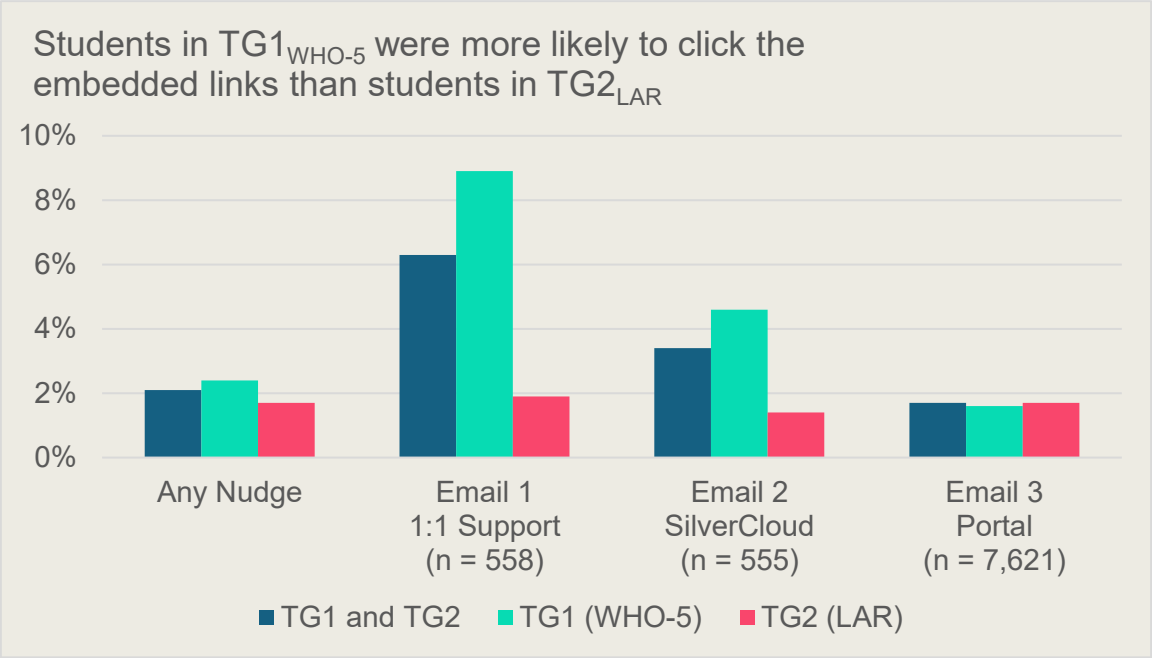


Figure 6: Email click-through rates for each treatment group and email type.

The open and click-through rates by levels of risk for students who were sent a nudge to one-to-one support or SilverCloud are presented in Table 23, ordered by open rate. Students who meet multiple definitions of risk (*n* = 169) have the highest open rate (44.4%) whilst students whose wellbeing is not at risk have the lowest (14.2%). Students who are at risk of a negative retention outcome as calculated by the learning analytics system have the lowest click-through rate (1.4%).

Table 23: Open and Click-through Rates by Risk Category

Risk Profile	Open Rate	Click-through Rate
At-risk for both WHO-5 and LAR (*n* = 169)	75 (44.4%)	8 (4.7%)
WHO-5 risk (*n* = 1,371)	501 (36.5%)	74 (5.4%)
LAR risk (*n* = 857)	246 (28.7%)	12 (1.4%)
Not at-risk (LAR) (*n* = 7,877)	1298 (16.5%)	170 (2.2%)
Not at-risk (WHO-5) (*n* = 7,363)	1043 (14.2%)	108 (1.5%)

Staff perspectives

In the focus group, staff commented on how the uptake in self-referrals did not show a clear spike that had been observed in previous periods when nudges had been sent out (JISC, 2024). However, participants did seem to differ in this perception as it was conflated with when cover was taking place during the half term period when self-referrals can show anomalous patterns.

> *"I think in the past it's been much more noticeable, and people seem to come through in a clump. I don't know what's different because it didn't feel like that. Maybe there was an increase overall in registrations after that nudge, but you feel like there's a spike usually."* (P6)

The participants did discuss why they felt a spike was not observed this time and they suggested 'email fatigue' may be a contributory factor as they sympathised with the high number of emails students receive

> *"So that begs the question, how long does a student take to process what they are being nudged for?"* (P3)

> *"Well, they're not the best at checking their emails."* (P5)

> *"That was my next point. Yeah. How many times have we sent an e-mail to a student and they ultimately say 'I've never read an email'".* (P3)

> *"I mean they get so much that I don't know how ... and I know that they're trying to do a lot about getting better at that, but I don't know how they find any e-mail at any point."* (P5)

> *"Which makes me question why were there such immediate responses in the past, but there definitely were."* (P2)

In relation to the effectiveness of the nudging intervention, staff drew comparison with a parallel campaign where all first-year students received a phone call to check in on how they were getting on. Staff suggested that, similar to nudging, the campaign had also not

shown the same demand on their resources as in previous years; though they did suggest that students may need time to reflect on these interventions before taking any immediate action. As such, this is important to consider when preparing staff for the days and weeks after proactive campaigns.

> *"I don't know how many of them come off the phone and then go on to [registration]. Because that's more likely to happen, isn't it? When you come off the phone, go think about it rather than the shock of some random person ringing you."* (P1)

There were both institutional- and individual-level variations that influence the extent to which the impact of the nudging interventions would be consistent over time and settings. Starting with an institutionally-led top-down perspective, the interpretation of the nudges by students was influenced by other communications the students received. Participants frequently drew parallels to attendance monitoring systems and how they reinforced a belief that any contact from support services was due to a student doing something wrong; and often questioned how such contact regarding mental health and wellbeing may be similarly received. Consequently, it is important to be cognisant of how changes in other systems and their methods of communications may impact the extent the interventions are perceived as *"supportive or punitive"* (P5).

> *"I had a student recently who was getting the emails because they weren't attending. And I said, well, what what's the e-mail address from? And he read the e-mail out to me and it was 'student conduct.' Which then made me initially go 'Oh, like that doesn't sound great.'"* (P6)

> *"Yeah, I guess in the mental health angle, it sounds like it's you saying 'There is something wrong with you' as if there's something going wrong. 'What's wrong?' as opposed to how 'Can we help you? You appear to not be attending.'"* (P5)

6.6.3 RQ3: Does proactively offering wellbeing support have an impact on students' propensity to seek help?

Focus group participants described their perceived determinants of students' engagement with support post-nudge. Prime examples were how perspectives of mental health by significant others and previous engagement at schools may create stigma and form barriers to help-seeking behaviour.

> *"I've had three students in the last two weeks have told me that their family don't believe in mental health. So that's the intrinsic barriers that they have in terms of stigma and trying to open up about it. And so anything that facilitates them seeking help is a direct challenge sometimes isn't it? 'Because you know me Dad says. I only went the counselling 'cause it was a skive From school.'"* (P6)

> *"My thought as I've worked in a school before coming here a year ago, when the school reached out to people, it was invariably because their attendance was poor, their behaviour was poor or there was a real chance they were going to be excluded or something like that. I wonder if there's anything that resonates from their school time where they've been through a difficult time and it's a self fulfilling prophecy. I'm going to get done because I'm not attending. I don't know if there's a link there."* (P4)

Participants also mentioned wider ways in which the nudging intervention may be effective beyond the immediate self-referrals. They highlighted how it may be reducing stigma at a wider level by encouraging discourse around mental health and promoting a 'whole university' approach to tackling mental health.

> *"One of things I have always advocated for in mental health is a free and easy discussion around mental ill health. And not being afraid to say and not being stigmatised. So if nudging helps that, that's great because you kind of want a mental health literate campus, don't you? Where people are OK to say 'I feel don't quite, You know,' and if nudging supports that kind of move towards that discourse, that's great."* (P1)

One of the hypothesised ways of measuring whether more students are seeking help to overcome challenges at university is to monitor the change of circumstances (CHoC) process. CHoC support initiations were therefore used as a proxy for students' propensity to seek help (see Table 24).

Table 24: CHoC-support initiations by treatment group

Group	Open Rate
Control group (*n* = 4,388)	65 (1.5%)
$TG1_{WHO-5}$ (*n* = 4,367)	52 (1.2%)
$TG2_{LAR}$ (*n* = 4,367)	55 (1.3%)
$TG1_{WHO-5}$ and $TG2_{LAR}$ (*n* = 8,734)	107 (1.2%)

Firstly, there was no significant difference between rates of CHoC-support initiations between nudged students (1.2%; n=107) and the control group (1.5%; n=65; $\chi2$ (1) =1.29, p = 0.26). Furthermore, there was no significant difference between the rate of CHoC-support initiations for students in $TG1_{WHO-5}$ compared to $TG2_{LAR}$ (1.2%, n= 52; $\chi2$(1) =0.04, p = 0.85).

6.6.4 RQ4: What are staff views on the different approaches to identifying students in need of support?

The results of the NOMAD-Lite survey, which was conducted at the beginning and end of the focus group are in Table 25. At the beginning of the session, every participant believed that participating in receiving referrals via nudging is a legitimate part of their role. Only two participants were able to distinguish how nudging differed from their usual way of working and, similarly, two participants felt staff had a shared understanding of the purpose of proactively nudging students.

Table 25: NOMAD-Lite Results both prior to and after the focus group. Changes in statement agreement at the beginning and end of the focus group are indicated with green (increase in agreement) or red (decrease in agreement)

	Prior to focus group				After focus group			
	Agree	Neither agree n or disagree	Disagree	Agree (%)	Agree	Neither agree n or disagree	Disagree	Agree (%)
I can see how nudging students towards services differs from usual ways of working	2	4	0	33%	4	2	0	67%
I believe that participating in receiving referrals via nudging is a legitimate part of m y role	6	0	0	100%	6	0	0	100%
There are key people who drive nudging inte rventions forward and get others involved	5	1	0	83%	5	1	0	83%
Management adequately support the nudging of students to services	5	1	0	83%	5	1	0	83%
I can see the potential value of nudging students for my work	4	2	0	67%	5	1	0	83%
I value the effects that nudging students has had on my work	4	1	1	67%	3	2	1	50%
Staff in this organisation have a shared understanding of the purpose of nudging students	2	3	1	33%	2	3	1	33%

At the end of the session, participants were encouraged to then revisit their initial responses and indicate if their responses had changed as a result of the discussion. One participant actually became less certain about whether they valued the effects that nudging had on their work (changing from 'Agree' to 'Neither Agree or Disagree'). Two participants felt that they better understood how to differentiate nudging from their usual way of working after the discussion. One staff member who was initially unsure about the value of nudging for their own work went on to indicate that they did see the value by the end of the focus group. There were no changes to any other responses.

General staff perspectives on data-informed nudging

Participants expressed that their role in the implementation of the intervention was primarily a 'readiness' for a potential influx of self-referrals. This was not seen as organising official cover for increased self-referrals but more creating an awareness so that practitioners felt prepared. There was division over how well this had been communicated as some expressed there had been ample notice whereas others felt unaware and were surprised of an influx. They also highlighted this had not necessarily been due to a lack of notification from within their team but felt like their manager had not been fully informed.

> *"We're aware of the nudging coming. And from our manager. But I think that was about it. In my view, we're aware that what was happening and then that we were probably going to plan for it. So that's the part that we needed to we planned for it knowing that it was going to happen maybe a possible spike. And we're involved in this."* (P4)

> *"There's been times where [manager], I think, hasn't known for some reason, yeah, but it's just gone out or been brought forward and you're thinking "What has happened to the registrations?!""* (P3)

> *"If you say management adequately support, I mean that's pretty ambiguous. But if you mean are we ready post nudge for an increase in registrations where, yeah, because again, looking at things like annual leave and who needs to be in and you know all that sort of operational nudge readiness was in place."* (P5)

When asked whether participants felt nudges went out at a good time of year they did feel that around half term was not well-timed as they could not attribute increased numbers due to the nudges and they felt that organisational cover was trickier during this period.

> *"I think yes, when you're knowing what's going to happen and things will be coming and you organise for it and you might put it in your diary."* (P2)

Normalisation Process Theory: Coherence

Most staff in the focus group clearly recognised how the nudging intervention was markedly different from previous methods of engagement that are primarily reactive. They recognised the uniqueness of nudging methods as being the individualism of the email and distinguished it from their current outreach activities.

> *"I think nudging is definitely different from a usual way of working as usual ways of working is we sit here and wait for them to come to us."* (P5)

> *"And any other service it would be that as there's not an awful lot of going about and trying to find people."* (P6)

> *"But we've done kind of outreach stuff, haven't we?"* (P5)

> *"But I guess the nudging is very specific to an individual, isn't it? Whereas if we go into a psychology lecture, say here we are, we're actively outreaching, but we're not walking in the lecture and going to joe blogs in the front row. 'Yeah, I'm here for you. I've seen your data.'* (P4)

An interesting observation was that, when staff participants were asked how this method of proactive intervention differs from other services, such as Disability services for providing Student Accessibility Plans (SAP), participants were able to differentiate based on the types of data used rather than focusing upon aspects more related to service models. They suggested that Disability services were reaching out on *"static data and not making live reviews"* (P3), perceiving the nudging intervention to be acting on live data that was indicative of risk rather than self-report alone.

> *"In the sense of a SAP, that person is saying 'I have this condition and I need help for this.' Whereas here what you're doing is you're saying 'We've noticed this and think you might need support for that.' So one's kind of a self-assessment. And the other one's an 'us' assessment."* (P5)

Normalisation Process Theory: Cognitive participation

Staff mentioned how proactive outreach was outside the scope of their service due to capacity issues with students already on their caseload. A fear of the researchers was how practitioners would perceive any increase in workload attributed to the intervention. However, staff recognised the value and necessity of early intervention if data suggested it was relevant. One participant drew a parallel in how practitioners already routinely act on data that is indicative of risk *without* a self-referral whilst others highlighted the potential moral obligations to act that could conflict with their role if they were responsible for the outreach themselves:

> "*On duty we do nudge people. There's people who are brought to our attention and then we actively go out and seek them so we don't always wait for them to come to us."*(P4)

> *"Potentially with dyslexia, if someone doesn't either recognise it or tell anyone for their three years at university, they could feasibly get to the end of the three years at university. I guess the difference is maybe there is a moral implication that if someone sat with severe mental health difficulties for three years, they might not get to the end of the three years? ...If it's telling you something risky you then have to do something about it."* (P5)

"I've worked in an assertive outreach team with the people who wouldn't engage with routine services, but there was always a kind of fine line between certainly trying to get someone into treatment and support and harassing them. I wouldn't want to be part of something, and I'm not saying it's harassment. But that kind of outreach in nudging and sort of saying 'perhaps you could…' It just wouldn't sit comfortably with me as a social worker." (P6)

"I think sometimes it feels a bit more comfortable, it being a bit more arm's length from us. I think it would absolutely change how students view us as well. I think. And that would be a danger. You've got to be very careful as we already tread a fine line." (P3)

Normalisation Process Theory: Collective action

Participants universally felt that it was important that they were not involved in the process of the dissemination of the nudges themselves and felt far more comfortable being the 'recipient' of the self-referrals that the nudges prompted. When asked whether they felt they were involved in the process one participant asked *"Do we have to be?"* (P5) and others expressed confusion as to whether there was an expectation that they should or could be more involved in the control of sending data-driven nudges. Participants recognised a delineation between involvement in the nudges and the services they provide, advocating for separation. They felt direct involvement would introduce ethical complications and more imperative to repeatedly act, even if their services are not demanded, for fear of being seen to abandon an obligation created by the original identifier of the risk.

"*Certainly we're not the key people who are driving the interventions forward*" (P1)

"*I'd be a little bit concerned that suddenly we'd have like a list of these 100 people come back to the idea of, like a moral obligation. Well, they're on our list and they've ignored us, but now 'cause we've reached out, are we now more culpable? Responsible. You know, 'cause, actually, we've reached out because we thought they needed it. They didn't respond. So then we left it. They could be like that on day one then the rest of the three years we're chasing this one person because we've identified them and we're scared to let it go.*"

Normalisation Process Theory: Reflexive monitoring

Participants clearly felt there was a benefit to nudging, suggesting the resulting referrals which came in post-intervention were worth the initial effort to analyse data and send nudges. However, they also recognised that there are ongoing capacity issues. While they spoke of how it inconvenienced them personally, the ethos of nudging was wholly supported as they recognised any method that increased the caseload with genuine student need was beneficial.

> *"I think we need to get as many people into the services we can, but I think there's a huge capacity issue. I think it's valuable. It would feel a shame, if it was a regular thing to then take it away 'cause you'd be thinking, well, there's some students who we would never get in touch with through other ways."* (P2)

Given, however, that staff felt strongly that the effort of identifying students for this activity was outside of their remit, it should therefore be noted that the capacity issue discussed here takes into account the resulting impact on caseload that proactive interventions have.

6.7 Summary of IPE findings

This section provides a summary of the key findings from the IPE and the caveats and limitations associated with this research.

Reach

Of the 24,575 enrolled students, 13,122 (54.1%) students consented to being part of the trial. Of those that received a nudge due to being deemed at risk, 4,367 (17.8%) were in $TG1_{WHO\text{-}5}$ and 4,367 (17.8%) were in $TG2_{LAR.}$

Female students were more likely than male students to consent to their wellbeing being measured using WHO-5, to consent to the use of analytics-driven support, and to consent to take part in the trial.

While white students were more likely to sign up for educational analytics support, students from BAME backgrounds had a higher comparative opt-in rate to wellbeing support and were more likely to consent to take part in the trial.

Students nudged based on their WHO-5 score were more likely to open emails signposting one-to-one support. Nearly half of the students who met both definitions of risk (n = 169) opened their nudge, the highest open rate by risk profile. Students who are at risk of a negative retention outcome as calculated by the learning analytics system had the lowest click-through rate (1.4%).

Effectiveness

Of the students who registered for one-to-one counselling, those who registered via the nudge email had similar levels of wellbeing (as measured by CORE-1) as those in the control group who self-referred.

Staff felt that the students nudged towards the support services warranted the support but suggested email fatigue may be an issue with students engaging with nudge-based support.

Implementation

The nudges were implemented according to the protocol, and 100% of emails were successfully delivered from the CRM system. However, staff reported that the nudge activity was not well timed for appropriate staff cover at half term. They also expressed that their role in the intervention was more 'readiness' for an influx in self-referrals rather than organising or managing the process.

Adoption

Staff recognised the value and necessity of early intervention if data suggested it was relevant but highlighted the potential moral obligation to act on data could conflict with their support role if they were responsible for identifying at-risk students themselves. Staff recognised the individualism of nudge-based support compared to other interventions.

CMH advisors were confident in their belief that participating in referrals via nudging is a legitimate part of their role as university.

6.7.1 Summary of caveats and limitations: Implementation and process evaluation

Whilst the present IPE has utilised several methods to explore and explain the implementation of a wellbeing nudge intervention there are several areas which need to be considered before adoption which include:

- **Student Voice**: whilst we have engaged with staff in this report to understand potential determinants of engagement with services, there is further work to do to engage with students to understand their perspectives on these nudges.
- **Maintenance:** This is a key aspect of the Glasgow et al. (1999) evaluation framework and describes the extent that the data-informed nudge intervention was (or could be) embedded and sustained as part of routine university practice. As the present intervention was a single delivery, actual maintenance beyond six months was not measure. However, a follow up study designed with this dimension in mind is recommended. Nevertheless, the study does consider adoption and the extent to which ongoing maintenance would be likely based on the barriers encountered in the present case.

7. Discussion

This study aimed to evaluate whether it is more effective to identify students for wellbeing support using their WHO-5 wellbeing score at enrolment or their LAR, generated by NU's learning analytics system.

Overall, we found no statistically significant impact of different methods of identifying at-risk students on students' likelihood of self-referring to support services. This is not straightforward to interpret, as it contains the bundled impact of both the targeting mechanism and an email intervention. All participants in $TG1_{WHO-5}$ and $TG2_{LAR}$ received an email of some description, so the lack of impact between the control and these conditions suggests that overall, the mixture of emails sent to the two treatment groups did not produce significantly higher outcomes than no email.

We found no statistically significant impact of the treatment on likelihood of self-referring to CMH services. Likewise, there was no practically meaningful difference between $TG1_{WHO-5}$ and $TG2_{LAR}$ and the control group with respect to self-referral to SilverCloud.

We found no statistically significant impact of the treatment on the other secondary outcomes of interest (LAR, likelihood of clicking on links embedded in the intervention email, and likelihood of withdrawing from university).

One notable finding was that students in $TG1_{WHO-5}$ had a directionally higher probability of engagement with the support links embedded in the email than $TG2_{LAR.}$ As discussed above, there are concerns about the robustness of the relationship due to variation in the significance and magnitude across models, along with the significance only being at the 10% threshold. However, this result suggests that perhaps the nudge emails may have been more effective in prompting a first step from those with low wellbeing identified via the WHO-5 score, compared to those with low LAR. However, we must take this finding, and any interpretation, with caution. Specifically, the difference we observe might also be explained by the fact that more students who fell below the WHO-5 threshold received the personalised nudge email as compared to those in $TG2_{LAR}$. Additionally, the effectiveness of WHO-5 over LAR was not evident in other outcomes. This could, however, also be because the other outcomes may take longer to materialise, in comparison to prompting the students to click a link embedded in an email.

To better understand the impact findings, we revisited the ToC of the intervention. The primary hypothesis of the trial focused on comparing the overall effectiveness of the identification method of at-risk students, via either the WHO-5 score or the LAR rating. The original intervention's ToC did not identify different mechanisms for the two groups; it stated that irrespective of which way they had been identified as at-risk, they

would act similarly when the emails had been delivered to them. This included assuming students would see the nudge emails as relevant and timely. However, a clearer and distinct theoretical mechanism of change for both risk-metrics should have been identified. The WHO-5 score identifies students based on their self-reported internal mental state. The theoretical mechanism could hypothesise that the student receiving the nudge email would feel validated as it would confirm what they had already recognised about their internal low wellbeing, and this would trigger help-seeking behaviour. In contrast, the LAR rating identifies students based on observable indicators such as low attendance. The ToC mechanism here was to provide students with a proactive warning. The students receiving the nudge email would see it as a warning that their attendance is being noticed and they are being notified to promptly act before the students' disengagement leads to more severe academic consequences. While these two distinct theoretical mechanisms were not explicitly stated in the intervention's ToC, the null impact show the gap between the theoretical expectations and the impact findings. The one-off, low intensity email was not sufficient to activate either mechanism fully.

Many factors identified in the IPE may have affected the null results and can help interpret the impact evaluation findings. First, the rate of consent processes directly impacted the reach of the intervention to the extent that the final number of students who took part in the trial was only 54.1% of the total student population. This limited reach likely contributed towards lack of intervention effectiveness. This might be because students who provided consent differed systematically from those who did not – for example, in motivation, engagement, or behavioural composition – which could dilute or mask the true impact of the intervention at the population level. For example, it is possible that students who did not provide consent were at higher risk and could have gained the most from the intervention. Hence, their exclusion from the trial may have limited our ability to observe the intervention's full impact. To maximise reach for future research and make it more representative of the true population, it is recommended that interventions consider whether it is necessary to streamline the consent processes beforehand.

Additionally, despite achieving a successful delivery of the nudge emails, overall engagement with the nudges was low. In fact, overall, only 17.7% of nudge emails were opened by students, and only 2.1% clicked through to the embedded support links. This aligns with prior evidence from studies involving similar-aged young people, which shows that the effectiveness of email nudges depends on students opening and actively engaging with the content (Taylor, 2021). Taken together, these findings indicate that email messaging may be an inherently limited channel for influencing young people in this context, and alternative forms of messaging should be considered in future trial designs

Additionally, different engagement patterns were observed between the two treatment groups. First, more students in $TG1_{WHO-5}$ opened and clicked the emails as compared to those in $TG2_{LAR}$. Next, data also showed that a larger proportion of students in $TG1_{WHO-5}$ opened emails signposting one-to-one support. Additionally, more students signposted to one-to-one wellbeing support also registered for it than those signposted to SilverCloud. This, therefore, suggests that the WHO-5 was effective in terms of a tool in identifying students who would respond to messaging than messages driven by academic analytics (LAR).

It is also noteworthy that generic emails, sent to those who did not fall below the thresholds of WHO-5 and LAR, signposting the standard wellbeing support had the lowest open rates overall. This supports the fact that targeting the right set of students is a crucial first step towards effectively delivering an intervention.

However, these immediate changes in behaviour, such as clicking on the embedded links, did not translate into further help-seeking behaviour such as self-referring to counselling support. One explanation for the low response rate was raised by staff who noted students are experiencing 'email fatigue' due to the sheer number of emails they receive and may need time to reflect before taking any immediate action.

In terms of implementation issues that may have affected the results, staff noted that the nudge intervention was not well-timed as they were sent out during half-term and there was not adequate staff cover to reach out and provide support to students or follow up. Additionally, staff noted there may be differences in the way students interpret the emails. Participants in the IPE shared that students receiving any contact from support services may have thought it was due to a student doing something wrong. Thus, it may be that some students perceived the intervention as intrusive or punitive which may have limited the effectiveness of the intervention.

The IPE found multiple factors that may explain the impact result:

- According to staff, email fatigue and feeling overwhelmed by the high number of university communications may have diluted the effectiveness of the nudge messages.
- According to staff the complex and multiple consent processes limited the reach of the intervention, reaching only just over half of the total student body (54.1%). This suggests streamlining the consent process may need to be explored in future interventions.
- Staff appreciated the intervention and found it potentially useful for students but acknowledged there are ongoing capacity issues and would not be able to meet the needs of students without increased resources.

- Some students may have found the emails not tailored to their specific needs, which may have limited their ability to act on the nudge email.
- Participants (staff and students) also shared issues with the timing of the nudge emails including delivery near a school break and overlap with other campaigns, which reduced staff preparedness and may have affected students' responsiveness.

Overall, we are not able to conclude that different approaches to identifying students who may benefit from an email nudging them towards one-to-one counselling or self-help resources are effective in helping those students re-engage with the university. However, the IPE suggests a number of ways that the present research could be built on or refined in future.

8. Roles and responsibilities

This evaluation was a collaboration between TASO, The Policy Institute at King's College London, and Northumbria University. Each organisation led a different aspect of the design, delivery, and evaluation of the wellbeing intervention.

TASO was the commissioning body and funder and advised on the design and management of this evaluation. It was also responsible for quality assuring its outputs, which was also supported by independent external peer reviews of this work. Rob Summers acted as the lead Research Manager. We would also like to acknowledge the contributions of Christoph Koerbitz, who supported the commissioning and management of the research, and Luke Arundel and Mikayla Boginsky for providing project and research support.

The impact evaluation was designed and delivered by The Policy Institute at King's College London. Susannah Hume acted as Principal Investigator, Beti Baraki as Co-Investigator and Project manager of the evaluation. Parnika Purwar provided quantitative analysis and research support, and Megan Liskey provided qualitative and research support.

NU delivered the intervention and led the design and delivery of the IPE. The IPE was supported by Dr Carly Foster and Dr James Newman.

Table 26: Roles and responsibilities

Organisation	Name	Role and responsibilities
TASO	Dr Rob Summers	Senior Research Manager: Project management
	Luke Arundel	Research Officer: Project support
	Mikayla Boginsky	Research Officer: Project support
	Christoph Koerbitz	Chief Research Officer: Commissioning oversight and management
King's College London	Dr Susannah Hume	Principal Investigator
	Beti Baraki	Co-investigator, project manager, qualitative lead
	Michael Sanders	Co-investigator, quantitative lead
	Parnika Purwar	Researcher, quantitative analysis and research support
	Megan Liskey	Researcher, research support
Northumbria University	Dr Carly Foster	Principal Investigator
	Dr James Newman	Co-Investigator & subject matter expert

9. Ethical considerations

9.1 Ethical approval

This research received ethical approval from Northumbria University Ethics Online System.

- Project No. 7742
- Ethical case reference: Newham 2024-7742-8041
- Date of ethical approval: 22/07/2024

The ethical approval for this project is conditional on adherence to the project scope and documents, covered within the application. Researchers must comply with Northumbria University's policies, procedures, guidance and standard operating procedures. These can be found on the website https://www.northumbria.ac.uk/research/ethics-integrity-and-trusted-research/.

9.2 Consent

For student data, consent to collect, process and analyse was captured digitally at enrolment using a participant information sheet which included how a student may withdraw from the project.

For staff data, consent to collect, process and analyse was captured prior to the focus group utilising a participant information sheet and consent form which was manually completed and signed.

9.3 Disclosure of harm

All focus groups were governed by Northumbria University's Ethical Governance in Research Policy. This policy sets out Northumbria University's approach to safeguarding and handling disclosures that may arise where we have serious concerns about the health or wellbeing of a research participant we come across while undertaking our project activities specifically;

> *"3.6 Limits to Confidentiality and Disclosure*
>
> *The confidentiality of information and research data should be respected within the limits of the law. Where applicable, consent procedures should make it clear that if something illegal, or potentially illegal, is discovered during a study, it may need to be disclosed to the proper authorities. Similarly, in order to safeguard the welfare of participants, researchers may need to share information with different authorities, should any risk, or history of harm or exploitation be disclosed during the research."*

9.4 Data protection

Northumbria University, Kings College London and TASO have a Data Sharing Agreement and agreed data protection protocols for this project. All policies and procedures in place to transfer, store, process, analyse, and dispose of data securely are in line with the 2018 Data Protection Act and UK GDPR requirements.

All student data for the evaluation has been de-identified before being securely transferred by Northumbria University to Kings College London for the evaluation's analysis. They received anonymous data with unique IDs only. The legal basis for the study's personal data processing is that the research is being conducted in the public interest, and/or is necessary for scientific and historical research purposes. All three parties were Joint Data Controllers for the duration of the evaluation for the anonymised data only. However, the University of Northumbria are owners of the anonymised data that is collected as business as usual on the Data Subjects, so may keep this data for their own purposes.

Focus group data was also recorded and stored on the research team's secure SharePoint folder at Northumbria University, with access restricted to only the research team and activity evaluation leads. The recording was deleted once it had been transcribed. Access to transcriptions were restricted to the same colleagues indicated above. The legal basis for processing this data is informed by participant consent.

As part of the ethics application, it was made clear the extensive protocols regarding data protection which needed to be adhered to safeguard participant privacy. In addition to an extension DPIA drafting and sign off process, the project also reviewed and published a new ethical code of practice for the use of student data for the purposes of educational analytics. Furthermore, the ethics process flagged that students should be made aware of the options available to them and that this should be done in a timely manner following consent therefore new wording was included in all enrolment confirmation letters which signposts wellbeing services.

10. References

Birch, H. A., Dodd, A., Francis, P., & Newham, J. (2026). Student insights on improving mental health and wellbeing communication at universities. *Perspectives: Policy and Practice in Higher Education*, *30*(1), 46–57. https://doi.org/10.1080/13603108.2025.2556189

Brooks, J., McCluskey, S., Turley, E., & King, N. (2015). The utility of template analysis in qualitative psychology research. *Qualitative research in psychology*, *12*(2), 202-222.

Clegg, H., Dodd, A., Francis, P., & Newham, J. J. (2026). Bringing student voice into mental health analytics: a preliminary exploration of communication preferences and intervention needs. *Perspectives: Policy and Practice in Higher Education*, *30*(1), 36–45. https://doi.org/10.1080/13603108.2025.2556184

Department for Education. (2024). HE Mental Health Implementation Taskforce—First Stage Report.

Finch, T. L., Rapley, T., Girling, M., Mair, F. S., Murray, E., Treweek, S., ... & May, C. R. (2013). Improving the normalization of complex interventions: measure development based on normalization process theory (NoMAD): study protocol. *Implementation science*, *8*(1), 43.

Foster, C. (2026). Student mental health profiling: using data for a whole university approach to mental health and wellbeing support. *Perspectives: Policy and Practice in Higher Education*, *30*(1), 4–15. https://doi.org/10.1080/13603108.2025.2556195

Glasgow, R.E., Vogt, T.M. and Boles, S.M. (1999) *Evaluating the public health impact of health promotion interventions: the RE-AIM framework,* Am J Public Health, 89:1322–7.

Hughes, G., Panjwani, M., Tulcidas, P., and Byrom, N. (2018) Student mental health: The role and experience of academics. Student Minds. Available at: https://www.studentminds.org.uk/uploads/3/7/8/4/3784584/180129_student_mental_health__the_role_and_experience_of_academics__student_minds_pdf.pdf [Accessed 12 June 2025].

Jones, C. S. and Bell, H. (2024) *Under increasing pressure in the wake of COVID-19: A systematic literature review of the factors affecting UK undergraduates with consideration of engagement, belonging, alienation and resilience.* Perspectives: Policy and Practice in Higher Education, 28(3), pp. 141-151. http://dx.doi.org/10.1080/13603108.2024.2317316

HEMHIT (Higher Education Mental Health Implementation Taskforce). 2024. Update report
2024. https://assets.publishing.service.gov.uk/media/6650ab78dc15efdddf1a8402/Student_Analytics_Update.pdf(open in a new window)

Keane, J. (2024). The Office for Students (OfS) mental health analytics project: an evaluation. *Accessed February 1*, 2025.

May, C., & Finch, T. (2009). Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology*, *43*(3), 535-554. https://doi.org/10.1177/0038038509103208

McLafferty, M., O'Neill, S., & Murray, E. K. (2024). Policy Brief Student Mental Health: College Student Mental Health and Wellbeing.

Morrish, A. 2023. *HE Providers' Policies and Practices to Support Student Mental Health*. London: Department for Education.

Newham, J. and Francis, P. (2018) *Suicide prevention and data analytics*. Office for Students. Available at: https://www.officeforstudents.org.uk/for-providers/equality-of-opportunity/effective-practice/suicide-prevention-and-data-analytics/ [Accessed 27 May 2025].

Newham, J., & Foster, C. (2025). Nudging university students to counselling and mental health services: staff perspectives on the implementation of a proactive well-being analytics intervention. *Perspectives: Policy and Practice in Higher Education*, 1-10.

Pereira, S., Early, N., Outar, L., Dimitrova, M., Walker, L., and Dzikiti C. (2020) *University student mental health survey 2020*: *A large scale study into the prevalence of student mental illness within UK universities.* The Insights Network. Available at: https://assets.website-files.com/602d05d13b303dec233e5ce3/60305923a557c3641f1a7808_Mental%20Health%20Report%202019%20(2020).pdf [Accessed 12 June 2025].

Peck, E., McCarthy, B., & Shaw, J. (2025). The Future of the Campus University: 10 Trends That Will Change Higher Education. HEPI Policy Note 64. *Higher Education Policy Institute*.

Pollard, E., Vanderlayden, J., Alexander, K., Borkin, H., & O'Mahony, J. (2021). Student mental health and wellbeing: insights from higher education providers and sector experts. *Community College Review*, *68*(14), 35-16.

Robertson, A., Mulcahy, E. and Baars, S. (2022) *What works to tackle mental health inequalities in higher education?* The Centre for Transforming Access and Student Outcomes in Higher Education (TASO). Available at: https://cdn.taso.org.uk/wp-content/uploads/2022-05-17_What-works-tackle-mental-health-inequalities-he_TASO.pdf [Accessed 27 May 2025].

Satherley, R.M., Newham, J., Cecil, E., Forman, J., Kelly, C., Wolfe, I. and Lingam, R., (2025). *Implementation of the children and young people's health partnership model of paediatric integrated care: a mixed-methods process evaluation.* Archives of Disease in Childhood. doi: 10.1136/archdischild-2024-327664

Taylor, P. (2021). *Service Learning and Youth Political Participation: A Mixed-Method Thesis*, Doctoral thesis, UCL (University College London).

Worsley, J., Pennington, A. and Corcoran, R. (2020) *What Interventions Improve College and University students' Mental Health and wellbeing? A review of review-level Evidence.* What Works Centre for Wellbeing. Available at: https://whatworkswellbeing.org/wp-content/uploads/2020/03/Student-mental-health-full-review.pdf [Accessed 27 May 2025].

11. Appendices

Appendix A: Impact table

This impact table is intended to summarise the results for the primary and secondary outcomes and communicate how confident we should be when making claims about the findings.

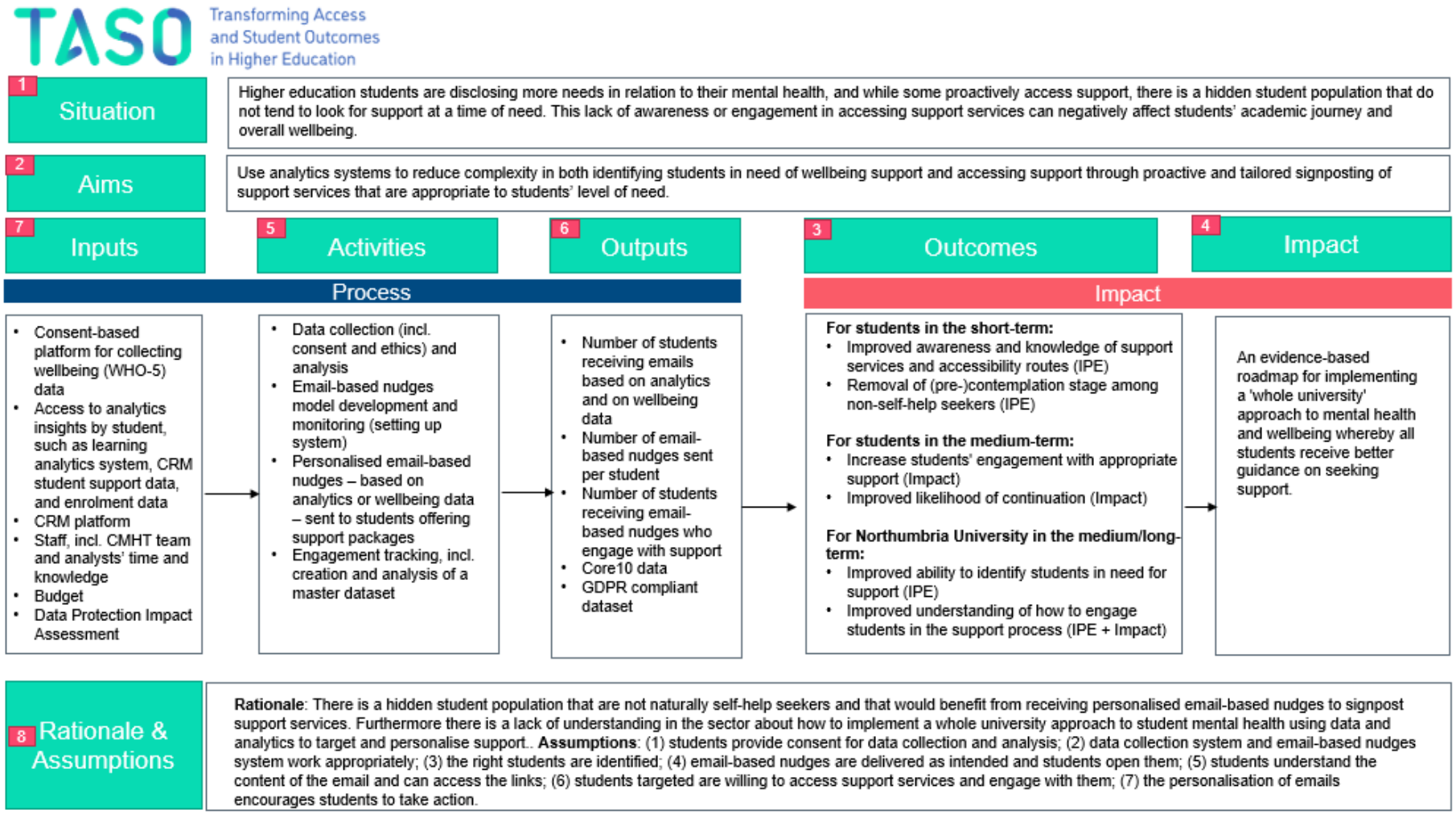
Table 27: Impact on outcomes summary

Outcome	Sample size	P Value	Effect	Estimated 'real world' effect	Evaluation security (1 = not at all secure 5 = very secure)
Self-referral to support services	6001	0.579	0.004 (-0.010, 0.019)	$TG1_{WHO-5}$ is 0.4 percentage points more likely to self-refer to support services as compared to control group but effect is statistically insignificant	5
	6001	0.621	-0.004 (-0.018, 0.011)	$TG2_{LAR}$ is 0.4 percentage points less likely to self-refer to support services as compared to control group but effect is statistically insignificant	5
LAR score	5845	0.767	0.002 (-0.006, 0.011)	$TG1_{WHO-5}$'s LAR score increases by 0.002 more than control group but effect is	5

Outcome	Sample size	P Value	Effect	Estimated 'real world' effect	Evaluation security (1 = not at all secure 5 = very secure)
				statistically insignificant	
	5845	0.801	0.002 (-0.007, 0.010)	$TG2_{LAR}$'s LAR score increases by 0.002 more than control group but effect is statistically insignificant	
Self-referral to CMH	6001	0.767	0.005 (-0.010, 0.019)	$TG1_{WHO-5}$ is 0.5 percentage points more likely to self-refer to CMH as compared to control group, but effect is statistically insignificant	5
	6001	0.801	-0.003 (-0.017, 0.011)	$TG2_{LAR}$ is 0.3 percentage points less likely to self-refer to CMH as compared to control group, but effect is statistically insignificant	5
Self-referral to SilverCloud self-help services	6001	0.983	0.000 (-0.004, 0.005)	There is no practically meaningful difference between $TG1_{WHO-5}$ and control group with respect	5

Outcome	Sample size	P Value	Effect	Estimated 'real world' effect	Evaluation security (1 = not at all secure 5 = very secure)
				to self-referral to SilverCloud	
	6001	0.801	-0.001 (-0.005, 0.004)	$TG2_{LAR}$ is 0.1 percentage points less likely to self-refer to SilverCloud as compared to control group, but effect is statistically insignificant	5
Withdrawal before Semester 2 commencement	6001	0.558	0.008 (-0.003, 0.019)	$TG1_{WHO-5}$ is 0.8 percentage points more likely to withdraw before Semester 2 commences as compared to control group but effect is statistically insignificant	5
	6001	0.749	0.007 (-0.004, 0.018)	$TG2_{LAR}$ is 0.7 percentage points more likely to withdraw before Semester 2 commences as compared to control group but effect is statistically insignificant	5

Appendix B: Theory of change illustration



Appendix C: Descriptive analysis

Table 28: Sample characteristics

	Full		$TG1_{WHO-5}$		$TG2_{LAR}$		Control	
Variable	n	%	n	%	n	%	n	%
Gender: Female	7,295	55.6	2,400	55.0	2,433	55.7	2,462	56.1
Gender: Male	5,827	44.4	1,967	45.0	1,934	44.3	1,926	43.9
Ethnicity: White	8,651	65.9	2,885	66.1	2,889	66.2	2,877	65.6
Ethnicity: BAME	4,212	32.1	1,395	31.9	1,396	32.0	1,421	32.4
Ethnicity: Unknown	259	2.0	87	2.0	82	1.9	90	2.1
Student: Continuing	6,940	52.9	2,310	52.9	2,320	53.1	2,310	52.6
Student: New	6,182	47.1	2,057	47.1	2,047	46.9	2,078	47.4
Faculty: Art, Design and Social Sciences (ADSS)	2,009	15.3	670	15.3	684	15.7	655	14.9
Faculty: Business and Law (BL)	3,633	27.7	1,226	28.1	1,202	27.5	1,205	27.5
Faculty: Electrical and Electronic Engineering (EE)	2,760	21.0	912	20.9	896	20.5	952	21.7
Faculty: Health and Life Sciences (HLS)	4,720	36.0	1,559	35.7	1,585	36.3	1,576	35.9
First in family: Yes	4,847	36.9	1,610	36.9	1,633	37.4	1,604	36.6
First in family: No	4,192	31.9	1,417	32.4	1,401	32.1	1,374	31.3
First in family: Unknown	4,083	31.1	1,340	30.7	1,333	30.5	1,410	32.1

Number of students: 13,122
Source: NU Management information

As stratification was carried out using WHO-5 and LAR score (see Section 4.2.3) there was a risk that randomisation would not result in well-balanced treatment and control groups across variables commonly used in stratification such as gender, ethnicity, and department. Broadly, the randomisation has resulted in treatment and control groups that are well balanced across gender, ethnicity, study stage, faculty, and family history of higher education. Additionally, the distribution between the treatment and control groups is similar. Gender distribution is relatively balanced, with both male and female participants distributed across the groups. Overall, the sample skewed slightly female (55.6 %). In terms of ethnicity, students are predominantly from white backgrounds (65.9%) while BAME students form about a third of the sample (32.1%). Ethnicity

appears well-balanced across treatment groups. On student status, 52.9% of the students in the sample are continuing students while 47.1% are new students. The distribution between control and treatment groups is well-balanced. Next, students span various departments with the largest proportion of students from the HLS department, standing at 36%. There are 27.7% from BL, 21% from EE and around 15% from ADSS. Finally, around 37% of the sample are the first in the family to go to university education; 31% did not share this information. The distribution between the control and treatment groups is similar.

In addition to the primary demographic variables, the distribution of the WHO-5 scores and the LAR scores are presented in Table 29 and Table 30 respectively. These scores were used for stratification to overcome the risk of confounding effects arising from differing baseline engagement and wellbeing metrics. Stratification has helped to ensure that WHO-5 and LAR scores are balanced across both treatment groups.

Table 29: Distribution by WHO-5 Score

	Full		$TG1_{WHO-5}$		$TG2_{LAR}$		Control	
Variable	n	%	N	%	n	%	n	%
0-50	2,086	15.9	692	15.8	679	15.5	715	16.3
50-100	11,036	84.1	3,675	84.2	3,688	84.5	3,673	83.7

Number of students: 13,122
Source: NU Management information

Table 30: Distribution by LAR

	Full		$TG1_{WHO-5}$		$TG2_{LAR}$		Control	
Variable	n	%	n	%	n	%	n	%
0-75	1,268	9.7	436	10	421	9.6	411	9.4
75-100	11,854	90.3	3,931	90	3,946	90.4	3,977	90.6

Number of students: 13,122
Source: NU Management information

Appendix D: Full regression tables

Regression 1: Full regression table of self-referral to support services on treatment assignment

	Model 1	Model 2	Model3	Model 4
(Intercept)	0.102 [0.076; 0.127] *	0.058 [0.048; 0.068] *	-2.129 (0.188) ***	0.102 [0.072; 0.132] *
Treatment1	0.004 [-0.010; 0.019]	0.004 [-0.011; 0.018]	0.072 (0.134)	
Treatment2	-0.004 [-0.018; 0.011]	-0.003 [-0.018; 0.011]	-0.071 (0.138)	-0.008 [-0.022; 0.007]
Faculty = BL	-0.043 [-0.065; -0.022] *		-0.739 (0.177) ***	-0.039 [-0.066; -0.012] *
Faculty = EE	-0.034 [-0.058; -0.011] *		-0.488 (0.186) **	-0.028 [-0.057; 0.001]
Faculty = HLS	-0.026 [-0.048; -0.005] *		-0.346 (0.141) *	-0.032 [-0.058; -0.007] *
Ethnicity = BAME	-0.001 [-0.017; 0.015]		-0.066 (0.160)	-0.004 [-0.025; 0.017]
Ethnicity = Unknown	-0.009 [-0.046; 0.028]		-0.201 (0.467)	-0.005 [-0.052; 0.041]
FiF = Unknown	-0.018 [-0.036; -0.001] *		-0.469 (0.185) *	-0.014 [-0.037; 0.010]
FiF = Yes	0.011 [-0.005; 0.027]		0.158 (0.125)	0.010 [-0.010; 0.029]
Gender = Female	0.008 [-0.006; 0.021]		0.149 (0.126)	0.012 [-0.005; 0.028]
Above 21 = Yes	-0.020 [-0.033; -0.008] *		-0.395 (0.121) **	-0.024 [-0.039; -0.008] *
Student status = New	-0.013 [-0.025; -0.001] *		-0.251 (0.113) *	-0.009 [-0.024; 0.006]
R^2	0.015	0.000		0.013
Adj. R^2	0.013	-0.000		0.011
Num. obs.	6001	6001	6001	4001
RMSE	0.233	0.234		0.233
AIC			2596.159	
BIC			2683.254	
Log Likelihood			-1285.079	
Deviance			2570.159	

**** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. 95% confidence intervals provided in brackets. Model 1 compares $TG2_{LAR}$ and $TG1_{WHO-5}$ to control group. Covariates in model 1 include gender, age group, ethnicity, first in family indicator, faculty and student stage. Regressions to conduct robustness checks in model 2 (no covariates) and 3 (logit model). Model 4 compares $TG2_{LAR}$ to $TG1_{WHO-5}$.*

Regression 2: Full Regression table of LAR on treatment assignment

	Model 1	Model 2	Model 3
(Intercept)	0.331 [0.295; 0.367] *	0.888 [0.880; 0.897] *	0.906 [0.884; 0.929] *
Treatment1	0.002 [-0.006; 0.011]	-0.001 [-0.013; 0.010]	
Treatment2	0.002 [-0.007; 0.010]	0.000 [-0.011; 0.012]	0.001 [-0.010; 0.012]
Faculty = BL	0.027 [0.015; 0.039] *		0.000 [-0.018; 0.019]
Faculty = EE	-0.004 [-0.017; 0.010]		-0.035 [-0.056; -0.013] *
Faculty = HLS	0.007 [-0.005; 0.019]		-0.034 [-0.052; -0.016] *
Ethnicity = BAME	0.015 [0.005; 0.025] *		0.021 [0.004; 0.038] *
Ethnicity = Unknown	0.016 [-0.017; 0.049]		-0.017 [-0.061; 0.027]
FiF = Unknown	0.020 [0.008; 0.031] *		-0.008 [-0.026; 0.011]
FiF = Yes	0.002 [-0.007; 0.010]		0.003 [-0.011; 0.017]
Gender = Female	0.000 [-0.008; 0.008]		0.020 [0.007; 0.033] *
Above 21 = Yes	-0.032 [-0.040; -0.024] *		-0.050 [-0.062; -0.038] *
Student status = New	0.023 [0.016; 0.031] *		0.014 [0.003; 0.026] *
R^2	0.426	0.000	0.035
Adj. R^2	0.424	-0.000	0.032
Num. obs.	5845	5845	3885
RMSE	0.137	0.181	0.178

**** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. 95% confidence intervals provided in the brackets. Model 1 compares $TG2_{LAR}$ and $TG1_{WHO-5}$ to control group. Covariates in model 1 include gender, age group, ethnicity, first in family indicator, faculty and student stage. Regressions to conduct robustness checks in model 2 (no covariates) and 3 (logit model). Model 4 compares $TG2_{LAR}$ to $TG1_{WHO-5.}$*

Regression 3: Full Regression table of self-referral to counselling services on treatment assignment

	Model 1	Model 2	Model 3	Model 4
(Intercept)	0.099 [0.074; 0.124] *	0.054 [0.045; 0.064] *	-2.147 (0.191) ***	0.100 [0.070; 0.130] *
Treatment1	0.005 [-0.010; 0.019]	0.004 [-0.010; 0.019]	0.085 (0.138)	-0.007 [-0.021; 0.007]
Treatment2	-0.003 [-0.017; 0.011]	-0.002 [-0.016; 0.012]	-0.054 (0.142)	
Faculty = BL	-0.042 [-0.063; -0.020] *		-0.728 (0.180) ***	-0.038 [-0.064; -0.012] *
Faculty = EE	-0.035 [-0.058; -0.012] *		-0.526 (0.190) **	-0.030 [-0.059; -0.002] *
Faculty = HLS	-0.028 [-0.048; -0.007] *		-0.384 (0.145) **	-0.031 [-0.056; -0.006] *
Ethnicity = BAME	0.001 [-0.015; 0.017]		-0.020 (0.163)	-0.002 [-0.023; 0.019]
Ethnicity = Unknown	-0.006 [-0.042; 0.031]		-0.130 (0.467)	-0.002 [-0.048; 0.044]
FiF = Unknown	-0.020 [-0.037; -0.003] *		-0.513 (0.190) **	-0.015 [-0.038; 0.009]
FiF = Yes	0.009 [-0.006; 0.025]		0.142 (0.128)	0.007 [-0.012; 0.026]
Gender = Female	0.005 [-0.008; 0.018]		0.100 (0.128)	0.007 [-0.009; 0.023]
Above 21 = Yes	-0.018 [-0.031; -0.006] *		-0.379 (0.124) **	-0.021 [-0.037; -0.006] *
Student status = New	-0.013 [-0.025; -0.001] *		-0.253 (0.116) *	-0.008 [-0.023; 0.007]
R^2	0.014	0.000		0.011
Adj. R^2	0.012	-0.000		0.009
Num. obs.	6001	6001	6001	4001
RMSE	0.227	0.228		0.228
AIC			2503.133	
BIC			2590.229	
Log Likelihood			-1238.567	
Deviance			2477.133	

**** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. 95% confidence intervals provided in the brackets. Model 1 compares $TG2_{LAR}$ and $TG1_{WHO-5}$ to control group. Covariates in model 1 include gender, age group, ethnicity, first in family indicator, faculty and student stage. Regressions to conduct robustness checks in model 2 (no covariates) and 3 (logit model). Model 4 compares $TG2_{LAR}$ to $TG1_{WHO-5}$.*

Regression 4: Full Regression table of self-referral to SilverCloud self-help services on treatment assignment

	Model 1	Model 2	Model3	Model 4
(Intercept)	0.004 [-0.003; 0.012]	0.006 [0.002; 0.009] *	-6.000 (0.740) ***	0.002 [-0.007; 0.011]
Treatment1	0.000 [-0.004; 0.005]	-0.000 [-0.005; 0.005]	0.014 (0.429)	
Treatment2	-0.001 [-0.005; 0.004]	-0.000 [-0.005; 0.004]	-0.116 (0.439)	-0.001 [-0.005; 0.004]
Faculty = BL	-0.006 [-0.012; 0.001]		-1.518 (0.796)	-0.004 [-0.012; 0.004]
Faculty = EE	-0.003 [-0.010; 0.005]		-0.226 (0.643)	0.000 [-0.009; 0.010]
Faculty = HLS	-0.001 [-0.008; 0.006]		-0.164 (0.423)	-0.003 [-0.011; 0.005]
Ethnicity = BAME	-0.001 [-0.006; 0.003]		-0.322 (0.558)	0.000 [-0.006; 0.006]
Ethnicity = Unknown	-0.005 [-0.008; -0.002] *		-14.161 (922.834)	-0.005 [-0.008; -0.001] *
FiF = Unknown	-0.002 [-0.006; 0.003]		-0.572 (0.650)	-0.003 [-0.007; 0.001]
FiF = Yes	0.001 [-0.004; 0.007]		0.199 (0.385)	0.005 [-0.002; 0.011]
Gender = Female	0.005 [0.002; 0.009] *		1.460 (0.569) *	0.007 [0.002; 0.012] *
Above 21 = Yes	0.000 [-0.004; 0.004]		0.017 (0.374)	0.000 [-0.004; 0.005]
Student status = New	0.002 [-0.002; 0.006]		0.351 (0.361)	0.001 [-0.003; 0.006]
R^2	0.004	0.000		0.005
Adj. R^2	0.002	-0.000		0.002
Num. obs.	6001	6001	6001	4001
RMSE	0.073	0.073		0.072
AIC			396.817	
BIC			483.913	
Log Likelihood			-185.408	
Deviance			370.817	

**** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. 95% confidence intervals provided in the brackets. Model 1 compares $TG2_{LAR}$ and $TG1_{WHO-5}$ to control group. Covariates in model 1 include gender, age group, ethnicity, first in family indicator, faculty and student stage. Regressions to conduct robustness checks in model 2 (no covariates) and 3 (logit model). Model 4 compares $TG2_{LAR}$ to $TG1_{WHO-5}$.*

Regression 5: Full Regression table of withdrawal before Semester 2 commencement on treatment assignment

	Model 1	Model 2	Model3	Model 4
(Intercept)	0.038 [0.021; 0.055] *	0.028 [0.021; 0.035] *	-3.322 (0.261) ***	0.045 [0.024; 0.067] *
Treatment1	0.008 [-0.003; 0.019]	0.009 [-0.002; 0.020]	0.270 (0.181)	
Treatment2	0.007 [-0.004; 0.018]	0.008 [-0.003; 0.019]	0.249 (0.182)	-0.001 [-0.012; 0.011]
Faculty=BL	-0.006 [-0.022; 0.009]		-0.186 (0.231)	0.002 [-0.017; 0.021]
Faculty=EE	-0.015 [-0.031; 0.001]		-0.506 (0.262)	-0.012 [-0.032; 0.007]
Faculty=HLS	0.000 [-0.014; 0.015]		0.002 (0.206)	0.007 [-0.011; 0.024]
Ethnicity=BAME	-0.006 [-0.017; 0.006]		-0.171 (0.212)	-0.006 [-0.021; 0.009]
Ethnicity=Unknown	-0.001 [-0.033; 0.032]		-0.003 (0.523)	-0.001 [-0.042; 0.039]
FiF=Unknown	-0.018 [-0.031; -0.005] *		-0.574 (0.232) *	-0.022 [-0.040; -0.005] *
FiF=Yes	-0.006 [-0.018; 0.006]		-0.170 (0.168)	-0.010 [-0.025; 0.005]
Gender=Female	-0.013 [-0.023; -0.003] *		-0.418 (0.157) **	-0.018 [-0.031; -0.004] *
Above 21=Yes	0.011 [0.000; 0.021] *		0.326 (0.150) *	0.009 [-0.004; 0.022]
Student status=New	0.015 [0.005; 0.024] *		0.459 (0.150) **	0.018 [0.006; 0.031] *
R^2	0.006	0.000		0.007
Adj. R^2	0.004	0.000		0.004
Num. obs.	6001	6001	6001	4001
RMSE	0.180	0.180		0.187
AIC			1750.671	
BIC			1837.767	
Log Likelihood			-862.336	
Deviance			1724.671	

**** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. 95% confidence intervals provided in the brackets. Model 1 compares $TG2_{LAR}$ and $TG1_{WHO-5}$ to control group. Covariates in model 1 include gender, age group, ethnicity, first in family indicator, faculty and student stage. Regressions to conduct robustness checks in model 2 (no covariates) and 3 (logit model). Model 4 compares $TG2_{LAR}$ to $TG1_{WHO-5}$.*

Regression 6: Full Regression table of subgroup analyses

	Gender	First in family	Faculty
(Intercept)	0.07 (0.05; 0.08) *	0.07 (0.05; 0.09) *	0.07 (0.04; 0.09) *
Gender = Female x $TG1_{WHO-5}$	-0.00 (-0.02; 0.01)		
Gender = Female x $TG2_{LAR}$	0.00 (-0.01; 0.02)		
FiF = Unknown x $TG1_{WHO-5}$		0.00 (-0.02; 0.02)	
FiF = Unknown x $TG2_{LAR}$		0.01 (-0.01; 0.03)	
FiF = Yes x $TG1_{WHO-5}$		0.01 (-0.01; 0.03)	
FiF = Yes x $TG2_{LAR}$		0.01 (-0.01; 0.03)	
Faculty = BL x $TG1_{WHO-5}$			0.00 (-0.03; 0.03)
Faculty = BL x $TG2_{LAR}$			0.01 (-0.02; 0.04)
Faculty = EE x $TG1_{WHO-5}$			0.00 (-0.03; 0.04)
Faculty = EE x $TG2_{LAR}$			-0.00 (-0.03; 0.03)
Faculty = HLS x $TG1_{WHO-5}$			-0.01 (-0.04; 0.02)
Faculty = HLS x $TG2_{LAR}$			-0.01 (-0.04; 0.03)
R^2	0.01	0.01	0.01
Adj. R^2	0.01	0.01	0.01
Num. obs.	13122	13122	13122
RMSE	0.20	0.20	0.20

**** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. 95% confidence intervals provided in the brackets. The models compare $TG2_{LAR}$ and $TG1_{WHO-5}$ to control group.*

Regression 7: Full Regression table of click-through status on treatment group

	Model 1	Model 2	Model3
(Intercept)	0.041 [0.020; 0.063] *	0.030 [0.023; 0.038] *	-3.393 (0.329) ***
Treatment2	-0.009 [-0.018; 0.001]	-0.008 [-0.018; 0.002]	-0.339 (0.202)
Faculty = BL	-0.024 [-0.042; -0.005] *		-0.743 (0.297) *
Faculty = EE	-0.014 [-0.036; 0.007]		-0.307 (0.311)
Faculty = HLS	-0.030 [-0.048; -0.012] *		-0.990 (0.266) ***
Ethnicity = BAME	0.001 [-0.012; 0.015]		0.077 (0.291)
Ethnicity = Unknown	-0.013 [-0.038; 0.012]		-0.769 (1.019)
FiF = Unknown	-0.018 [-0.033; -0.003] *		-0.765 (0.331) *
FiF = Yes	-0.005 [-0.018; 0.008]		-0.161 (0.229)
Gender = Female	0.014 [0.003; 0.025] *		0.597 (0.235) *
Above 21 = Yes	0.015 [0.004; 0.026] *		0.597 (0.209) **
Student status = New	0.003 [-0.008; 0.013]		0.128 (0.203)
R^2	0.010	0.001	
Adj. R^2	0.007	0.000	
Num. obs.	4001	4001	4001
RMSE	0.159	0.160	
AIC			958.087
BIC			1033.619
Log Likelihood			-467.044
Deviance			934.087

**** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. 95% confidence intervals provided in the brackets. Model 1 compares $TG2_{LAR}$ to $TG1_{WHO-5}$. Covariates in model 1 include gender, age group, ethnicity, first in family indicator, faculty and student stage. 1. Regressions to conduct robustness checks in model 2 (no covariates) and model 3 (logit regression)*

Appendix Dii: Additional regression analysis

Table 31: Exploring how first in family to attend university and student status are correlated with email clicks

Subgroup	Treatment Group	Interaction Estimate	Standard Error	p value
Student status = Continuing (ref: New)	1	-0.003	0.008	0.699
	2	0.001	0.007	0.906
First in Family = Yes (ref: No)	1	-0.010	0.010	0.338
	2	-0.002	0.009	0.779
First in Family = Yes (ref: Unknown)	1	-0.018+	0.009	0.054
	2	-0.014+	0.008	0.080

Table 32: Exploring the sensitivity of the threshold deployed by testing the interaction of the treatment with the baseline LAR and WHO5 score on post-intervention LAR

	Analytical Sample		Full Sample	
	Model 1	Model 2	Model 3	Model 4
(Intercept)	0.075 [0.022; 0.128]*	0.159 [0.119; 0.200]*	0.089 [0.044; 0.133]*	0.142 [0.110; 0.175]*
Treatment1	-0.021 [-0.087; 0.044]	0.020 [-0.031; 0.070]	-0.022 [-0.080; 0.036]	0.026 [-0.017; 0.070]
Treatment2	-0.005 [-0.066; 0.057]	-0.022 [-0.070; 0.026]	-0.008 [-0.063; 0.047]	-0.014 [-0.055; 0.028]
TG1 x Baseline LAR	0.031 [-0.050; 0.113]		0.027 [-0.038; 0.092]	
TG2 x Baseline LAR	0.001 [-0.075; 0.078]		0.005 [-0.056; 0.067]	
TG1 x Baseline WHO-5 score		-0.000 [-0.001; 0.000]		-0.000 [-0.001; 0.000]

	Analytical Sample		Full Sample	
	Model 1	Model 2	Model 3	Model 4
TG2 x Baseline WHO-5 score		0.000 [-0.000; 0.001]		0.000 [-0.000; 0.001]
Baseline LAR	0.030 [-0.026; 0.086]		-0.023 [-0.069; 0.023]	
Faculty=BL	-0.042 [-0.064; -0.020] *	-0.034 [-0.055; -0.012] *	-0.040 [-0.053; -0.028] *	-0.032 [-0.045; -0.019] *
Faculty=EE	-0.033 [-0.056; -0.009] *	-0.029 [-0.052; -0.005] *	-0.031 [-0.045; -0.017] *	-0.025 [-0.039; -0.011] *
Faculty=HLS	-0.024 [-0.046; -0.003] *	-0.020 [-0.041; 0.002]	-0.024 [-0.037; -0.011] *	-0.017 [-0.030; -0.004] *
Ethnicity=BAME	-0.001 [-0.017; 0.015]	0.004 [-0.012; 0.020]	-0.003 [-0.012; 0.007]	0.002 [-0.008; 0.011]
Ethnicity=Unknown	-0.008 [-0.044; 0.029]	-0.004 [-0.041; 0.032]	0.001 [-0.023; 0.025]	0.004 [-0.020; 0.028]
FiF=Unknown	-0.018 [-0.035; -0.000] *	-0.010 [-0.027; 0.008]	-0.009 [-0.020; 0.001]	-0.002 [-0.013; 0.008]
FiF=Yes	0.010 [-0.006; 0.027]	0.012 [-0.004; 0.028]	0.010 [0.001; 0.019] *	0.012 [0.002; 0.021] *
Gender=Female	0.006 [-0.007; 0.020]	0.000 [-0.013; 0.014]	0.011 [0.004; 0.019] *	0.006 [-0.001; 0.014]
Above 21=Yes	-0.019 [-0.031; -0.006] *	-0.013 [-0.026; -0.000] *	-0.005 [-0.013; 0.003]	-0.003 [-0.011; 0.004]
Student status=New	-0.012 [-0.025; -0.000] *	0.002 [-0.012; 0.015]	-0.006 [-0.013; 0.001]	-0.001 [-0.008; 0.006]

	Analytical Sample		Full Sample	
	Model 1	Model 2	Model 3	Model 4
R^2	0.016	0.026	0.010	0.024
Adj. R^2	0.013	0.023	0.009	0.023
Num. obs.	6001	6001	13122	13122
RMSE	0.232	0.231	0.201	0.199

**** p < 0.001; ** p < 0.01; * p < 0.05; + p<0.1. 95% confidence intervals provided in the brackets. Model 1 compares $TG2_{LAR}$ to $TG1_{WHO-5}$. Covariates in model 1 include gender, age group, ethnicity, first in family indicator, faculty and student stage. 1. Regressions to conduct robustness checks in model 2 (no covariates) and model 3 (logit regression).*

Appendix E: Data collection tools for IPE

Focus group discussion guide (NOMAD-Lite)

Seven statements were provided on a printed sheet prior to the focus group commencing so that participants could consider their views and respond based on a 3-point scale: Agree, Disagree, Neither Agree nor Disagree. A description of the term 'nudging' was also given for those who were unfamiliar with the term in this context: "'Nudging' in this scenario refers to using student data to identify students potentially at-risk of mental health difficulties and sending messages that direct them to wellbeing services that might be beneficial to them."

Question	Agree	Neither agree/ disagree	Disagree
1. I can see how nudging students towards services differs from usual ways of working			
2. I believe that participating in receiving referrals via nudging is a legitimate part of my role			
3. There are key people who drive nudging interventions forward and get others involved			
4. Management adequately support the nudging of students to services			
5. I can see the potential value of nudging students for my work			
6. I value the effects that nudging students has had on my work			
7. Staff in this organisation have a shared understanding of the purpose of nudging students			

Appendix Eii: Participant Information Sheet (Staff)

Title: Exploring Student Mental Health and Wellbeing (SMHW) staff's views of signposting services

Participant Information Sheet

You are being invited to take part in this research study. Before you decide if you would like to take part, it is important that you read this document so you understand why the study is being carried out and what it will involve. Reading this document, discussing it with others, or asking any questions might help you decide whether you would like to take part.

What is the Purpose of the Study?

As part of a wider project, we have trialled using data to identify students who may be struggling with mental health difficulties and sent targeted messages accordingly. We want to speak with staff within the Student and Mental Health Team to hear how this intervention may have impacted your ways of working. We want to understand staff's concerns, barriers, and benefits of this intervention.

Why have I been invited?

We are recruiting staff within the Student Mental Health and Wellbeing Team at Northumbria University.

What will happen if I take part?

We would like you to take part in an group discussion that typically takes 30-60 minutes. You will be asked to complete a brief questionnaire asking about ways the intervention may have impacted your service and your understanding of the intervention. This will be used to inform a discussion with your colleagues. There are no right or wrong answers, and we want to hear a range of opinions.

The focus group will take place at a time when you are not expected to have any appointments, which we have coordinated with the head of the service. With your consent we would like to record the audio of the discussion (not video). The recording will be written up so that we have an accurate record of the discussion to refer to afterwards, but all identifiable information will be removed from the transcript. We will also take handwritten notes in case of any problems with the recording.

What do I do if I want to take part?

If you would like to help us by taking part in this group discussion, or would like further information, please contact James Newham via email (james.newham@northumbria.ac.uk) or alternatively speak to Ellen Smith, Head of Student Mental Health and Wellbeing (ellen2.smith@northumbria.ac.uk)

Do I have to take part?

No. It is up to you whether you would like to take part in the study. I am giving you this information sheet to help you make that decision. If you do decide to take part, remember that you can stop being involved in the study whenever you choose, without telling me why.

What are the possible disadvantages of taking part?

This study will take approximately 60 minutes for people to complete. We recognise that this might seem like some time. However, people often find completing such studies interesting as it is an opportunity to voice your concerns and thoughts on the new intervention within the university.

What are the possible benefits of taking part?

There are no direct benefits to taking part in this project, however, the information you provide us will help us understand how we can support staff in our guideline recommendations.

Will my taking part in this study be kept confidential and anonymous?

Only the research team will have access to the data provided. Your personal name will not be used in any analysis. Any quotes will be anonymised so you cannot be identified. The interview recording will be destroyed after transcription is complete. After data collection, all information will be stored on an encrypted file location that only the research team has access to. All identifiable data will be destroyed after the investigation has been completed.

What categories of personal data will be collected and processed in this study?

The only personally identifiable information and data gathered during this research is your name and contact details to arrange the meeting. This information will be destroyed as soon as it is no longer needed (e.g. email addresses used to keep in contact with you will be destroyed as soon as they are no longer required). Consent forms with personal details will be destroyed at the conclusion of the project.

What will happen to the results of the study? Could personal data collected be used in future research?

The general findings might be reported in a scientific journal or presented at a research conference, however the data will be anonymized and you or the data you have provided will not be personally identifiable. We can provide you with a summary of the findings from the study if you email the researcher at the address listed below.

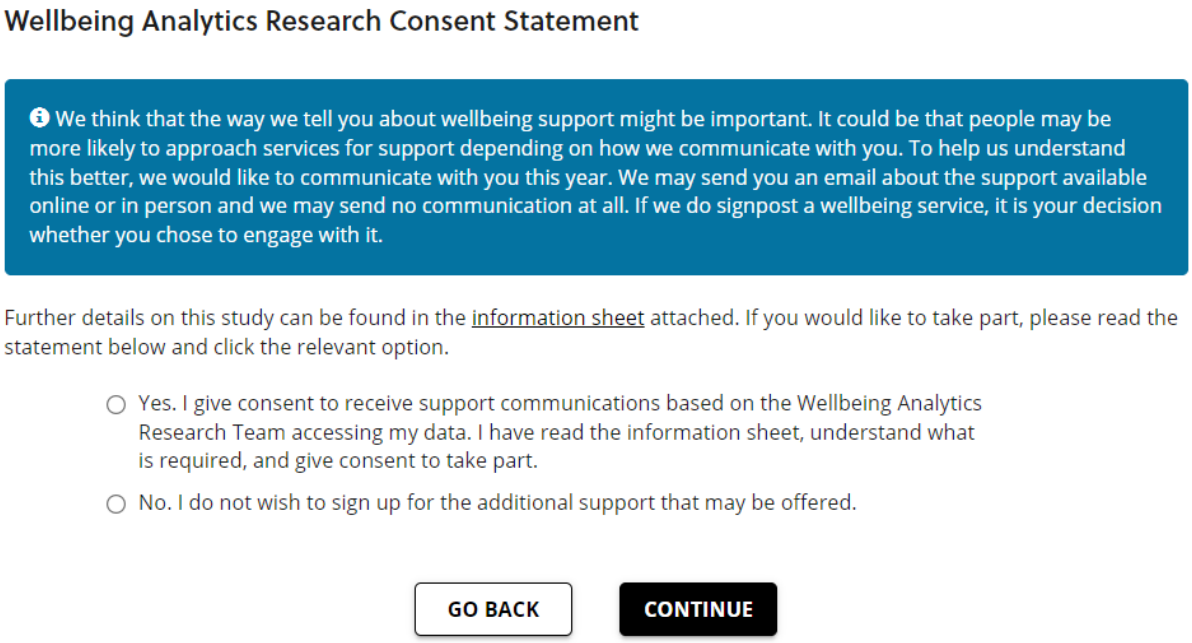
Who is organising and funding the study?

The research is funded by TASO (Centre for Transforming Access and Student Outcomes in Higher Education).

Who has reviewed this study?

The research project, submission reference 7742 has been approved in Northumbria University's Ethics Online system. It has been reviewed to safeguard your interests and have granted approval to conduct the study. If you have any concerns or worries concerning this research or if you wish to register a complaint, please direct it to the Department of Psychology Ethics Chair at emma.barkus@northumbria.ac.uk.

Appendix Eiii: Enrolment Screen and Participant Information Sheet (Students)



Information sheet for Northumbria University Analytics Initiative

Version 1 - 19th June 2024

The purpose of this information sheet is to provide you with relevant information so that you can give your informed consent. Please read this document carefully and raise any issues that you do not understand with the investigator.

What is the purpose of the initiative?

Northumbria University processes a wide variety of data to identify when a student may be struggling academically or with their mental health. This includes information you provide the university, attendance data, IT usage, and academic record. By doing so, we can intervene earlier and signpost you to services that may offer support.

As part of this initiative, we think that it is important we understand (i) what data may be important in identifying students who are struggling, and (ii) the best way to communicate with you. As you have consented to either learning or wellbeing analytics as part of the enrolment task, we would like to send you a follow up message. The content of this message will be randomised so that you hear about alternative support options, or we may send no communication at all.

Why have I been selected to take part and what are the exclusion criteria?

We are asking every student at Northumbria University to take part. There are no exclusion criteria.

What will I have to do?

You will not have to do anything other than provide your consent. This will allow our Student Life and Wellbeing team to send you a randomised message tailored to information we have on you, and monitor which students respond to the messages sent. It will be up to you whether you would like to follow up with the support we describe in the email.

What is the impact on me?

You may receive additional information support that may be available. All students consenting will receive an email of available services.

Will my participation involve any physical or psychological discomfort, or embarrassment?

You will not experience any physical discomfort during this study.

How will confidentiality be assured and who will have access to the information that I provide?

Only the Student Life and Wellbeing team with an embedded researcher will have access to the data provided. To link your responses to data that is routinely collected, the research team will generate a unique code from your student ID number to prevent identifying you personally in the data analysis. Your personal name will not be used in any analysis. After data collection, all information will be stored on an encrypted file location that only the research team has access to. All data will be destroyed 3 years after the investigation has been completed.

Will I receive any financial rewards / travel expenses for taking part?

As this is an online survey, no financial rewards or travel expenses will be given.

How can I withdraw from the project?

Please contact the researcher before October 2024 to ensure that you are not sent any messages (james.newham@northumbria.ac.uk). You can also change your opt-in preferences for either learning or wellbeing analytics at any time by updating your details via My Info in the Student Portal.

Data Protection and Confidentiality

Any personally identifiable information and data gathered during this research is subject to and will be stored in line with EU General Data Protection Regulation (GDPR) and the UK Data Protection Act (2018). The legal basis for the study's personal data processing is that the research is being conducted in the public interest, and/or is necessary for scientific and historical research purposes. For more information on how research data is processed by Northumbria University, and your rights under the GDPR, please see our Research Participant Privacy Notice

If I require further information who should I contact and how?

For further information, please contact the lead researcher via email (james.newham@northumbria.ac.uk). If you have any concerns or worries concerning this research or if you wish to register a complaint, please direct it to the Department of Psychology Ethics Chair at nick.neave@northumbria.ac.uk. This study and its protocol have received full ethical approval from the Department of Psychology Ethics Committee. If you require confirmation of this please contact the Chair of this Committee (Nick Neave), stating the title of the research project and the name of the researcher.

Appendix F: Intervention design details

Pre-randomisation: All consenting students will complete the wellbeing questionnaire at enrolment which runs from July-October. The data is stored in the student record system (SITS) and needs to be downloaded and a formula needs to be applied to the raw scores to generate a total score. At the moment anyone who scores <50 would be classified as low wellbeing and would be included in the trial. Students are then randomised to one of three groups (treatment group 1, treatment group 2, or control group)

$TG1_{WHO-5}$**:** Students in the 'Low' and 'Very Low' wellbeing groups (as assessed using WHO-5 survey) will be further randomised to receive a nudge promoting Counselling or SilverCloud. Students scoring 'Average' or 'High' will receive a generic email with no specific 'call to action' other than to review the wellbeing page on the student portal.

$TG2_{LAR}$**:** Northumbria University uses a learning analytics dashboard (Illume, created by Civitas) that draws data from a range of institutional systems to track how engaged a student is with their learning. These data include virtual learning environment (VLE) activity, attendance monitoring, submissions, grade performance, building access. The dashboard generates a score for predicted likelihood of a student having a positive retention outcome (repeat, progress or complete) the following academic year (0-100%). Those students below 75% receive nudges regularly from the student success team however this support is not aligned to wellbeing specifically and is more generic on course support.

Students indicated 'at risk' of retention (<75%) will be further randomised and receive either a nudge promoting self-referral to Counselling or SilverCloud. Students not 'at risk' will receive a generic email with no specific 'call to action' other than to review the wellbeing page on the student portal.

Control: These students will not receive a nudge regardless of whether their WHO-5 score or learning analytics score is low or very low.

Nudge 1 Email

Subject: Signposting you to Wellbeing Support via 1:1 Support

Hi there!

I'm Ellen, Head of the Student Wellbeing and Mental Health team at Northumbria University. I wanted to let you know about a great resource available to you.

We offer short-term, one-to-one therapeutic interventions designed to support you throughout your time at university. By having a conversation, you will be able to explore your personal thoughts, feelings, and experiences in a safe space. You will not have to talk about anything you do not wish to discuss.

To get started, complete this quick registration form. We'll review it and send you an email with a link to book a 15-minute telephone triage appointment with one of our Practitioners. This helps us understand your needs and ensure you get the right support. After your triage, you can choose to attend a 45-minute Wellbeing Plan Appointment for more personalised, one-to-one support.

If you have any questions or need help, just let us know!

Take care,

The Northumbria University Mental Health and Wellbeing Team

P.S. Check out this article on your portal to see what it's all about.

Nudge 2 Email

Subject: Signposting you to Wellbeing Support via SilverCloud

Hi there!

I'm Ellen, Head of the Student Wellbeing and Mental Health Team at Northumbria. I want to share a useful resource with you...

As a Northumbria student, you have access to SilverCloud, an online therapy platform. It's a free, online wellbeing programme that you can use anytime, 24/7. You'll find support for things like tackling anxiety and depression, practicing mindfulness, building resilience, enhancing sleep, and reducing stress.

Silvercloud uses proven methods to help people manage challenges by encouraging them to change the way they think and behave. It's particularly suited for people who don't like the idea of face-to-face appointments; you work through resources at your own pace, *without* a practitioner.

To get started, log in to SilverCloud here with your Northumbria email address.

If you have any questions or need help, just let us know!

Take care,

The Northumbria University Mental Health and Wellbeing Team

P.S. Check out this quick intro to SilverCloud here to see what it's all about.

Nudge 3 Generic Email

Subject: Signposting you to Wellbeing Support on your Student Portal

Hi there!

I'm Ellen, Head of the Student Wellbeing and Mental Health team at Northumbria. I just wanted to point you towards some useful resources available on our student portal. There are a lot of helpful articles, but here are a couple of key ones to get you started:

An overview of mental health and wellbeing support at Northumbria

Look after your wellbeing and access self-care resources to start or keep feeling good

Weekly Wellbeing Resources Drop-in

Feel free to explore these, and other links, on your portal!

Take care,

The Northumbria University Mental Health and Wellbeing Team

Appendix G: Code for randomisation / analysis

Appendix G(i) Randomisation code for top level randomisation

```
library(dplyr)

data <- read_excel("(FILEPATH)/(FILENAME).xlsx") #change URL

#Create strata

data<-data(order(data$Rand_code),) #order data by ID

data$strata<-ifelse(data$LA_OCT>=median(data$LA_OCT) &
data$WHO5>=median(data$WHO5),1,

                    ifelse(data$WHO5>=median(data$WHO5),2,
                    ifelse(data$LA_OCT>=median(data$LA_OCT),3,4
                    )))

set.seed(123)

prop1<-0.333 #set proportion assigned to arm1

prop2<-0.333 #set proportion assigned to arm2

prop3<- 1-(prop1+prop2) #implied proportion assigned to
control

data<-data %>% group_by(strata) %>% mutate(treat={

    n <- n()

    allocation1 <- floor(n*prop1) #proportion of stratum
allocated to arm1

    allocation2 <- floor(n*prop2) #proportion of stratum
allocated to arm2

    assign <- c(rep(1, allocation1), rep(2, allocation2),
rep(0,n- (allocation1+allocation2))) #vector of 1s (treat) and
0s (control)

    sample(assign) #randomise order of assignment within each
stratum

  }) %>% ungroup()

table(data$strata,data$treat) #check allocation
```

Appendix G(ii): Randomisation code for second level randomisation

```
setwd("FILEPATH ")


library(dplyr)

data <- read.csv("TASO_DATA.csv")


#Create strata

data<-data(order(data$Rand_code),) #order data by ID

data$strata<-ifelse(data$LA_OCT>=median(data$LA_OCT) &
data$WHO5>=median(data$WHO5),1,

                    ifelse(data$WHO5>=median(data$WHO5),2,


ifelse(data$LA_OCT>=median(data$LA_OCT),3,4)))


set.seed(123)

prop1<-0.333 #set proportion assigned to arm1

prop2<-0.333 #set proportion assigned to arm2

prop3<-1-(prop1+prop2) #implied proportion assigned to control

data<-data %>% group_by(strata) %>% mutate(treat={

    n <- n()

    allocation1 <- floor(n*prop1) #proportion of stratum
allocated to arm1

    allocation2 <- floor(n*prop2) #proportion of stratum
allocated to arm2

    assign <- c(rep(1, allocation1), rep(2, allocation2),
rep(0,n- (allocation1+allocation2))) #vector of 1s (treat) and
0s (control)
```

```
    sample(assign) #randomise order of assignment within each
stratum

  }) %>% ungroup()


table(data$strata,data$treat) #check allocation


## Creating the risk list in each treatment group and randomly
assigning intervention which is 50/50


prop4<- 0.5

prop5<-0.5


#At Risk in Treatment Group 1 (WHO5 GROUP)


tg1<- subset(data,data$treat==1 & data$WHO5<50)


tg1$strata<-ifelse(tg1$WHO5>=median(tg1$WHO5),1,2)


tg1<-tg1 %>% group_by(strata) %>% mutate(nudgetreat={

  n <- n()

  allocation1 <- floor(n*prop4) #proportion of stratum
allocated to arm1

  allocation2 <- floor(n*prop5) #proportion of stratum
allocated to arm2

  assign <- c(rep(1, allocation1),  rep(0,n- (allocation1)))
#vector of 1s (treat) and 0s (control)

  sample(assign) #randomise order of assignment within each
stratum
```

```
}) %>% ungroup()



#Treatment Group 2 (LA GROUP)


tg2<-subset(data,data$treat==2& data$LA_OCT<0.75)

tg2$strata<-ifelse(tg2$LA_OCT>=median(tg2$LA_OCT),1,2)


tg2<-tg2 %>% group_by(strata) %>% mutate(nudgetreat={

  n <- n()

  allocation1 <- floor(n*prop4) #proportion of stratum
allocated to arm1

  allocation2 <- floor(n*prop5) #proportion of stratum
allocated to arm2

  assign <- c(rep(1, allocation1),  rep(0,n- (allocation1)))
#vector of 1s (treat) and 0s (control)

  sample(assign) #randomise order of assignment within each
stratum

}) %>% ungroup()



riskdata<-rbind(tg1,tg2)


#create nudgelist by merging riskdata with those not at risk
in TG1 or TG2 but who are receiving an email


noriskdata<-subset(data, data$treat !=0)
```

```
noriskdata<-noriskdata %>%
  anti_join(riskdata,by = "Rand_code")
noriskdata$nudgetreat<-2


nudgelist <-rbind(noriskdata,riskdata)


nudgelist <- nudgelist %>%
  mutate(treat = case_when(
    treat == 1 ~ "WHO5",
    treat == 2 ~ "LA",
    TRUE ~ as.character(treat)  # Keep other values as they
are
  ))


nudgelist <- nudgelist %>%
  mutate(nudgetreat = case_when(
    nudgetreat == 0 ~ " 1:1 ",
    nudgetreat == 1 ~ "SilverCloud",
    nudgetreat == 2 ~ "KBA",
    TRUE ~ as.character(nudgetreat)  # Keep other values as
they are
  ))



write.csv(nudgelist,"nudgelist.csv", row.names = FALSE)
```

```
table(riskdata$strata,riskdata$treat) #check allocation


table(nudgelist$nudgetreat,nudgelist$treat) #check allocation


#when this code was run at 13:08 on 16/10/2024 the
intervention list looked as follows:-

#            LA TG   WHO5 TG

#1:1          205     336

#KBA emails   3934    3672

#SilverCloud  204     335
```

Appendix H: Randomisation code for second level randomisation

```
setwd("FILEPATH ")


library(dplyr)

data <- read.csv("TASO_DATA.csv")


#Create strata

data<-data[order(data$Rand_code),] #order data by ID

data$strata<-ifelse(data$LA_OCT>=median(data$LA_OCT) &
data$WHO5>=median(data$WHO5),1,

                    ifelse(data$WHO5>=median(data$WHO5),2,


ifelse(data$LA_OCT>=median(data$LA_OCT),3,4)))


set.seed(123)

prop1<-0.333 #set proportion assigned to arm1

prop2<-0.333 #set proportion assigned to arm2

prop3<-1-(prop1+prop2) #implied proportion assigned to control

data<-data %>% group_by(strata) %>% mutate(treat={

    n <- n()

    allocation1 <- floor(n*prop1) #proportion of stratum
allocated to arm1

    allocation2 <- floor(n*prop2) #proportion of stratum
allocated to arm2

    assign <- c(rep(1, allocation1), rep(2, allocation2),
rep(0,n- (allocation1+allocation2))) #vector of 1s (treat) and
0s (control)
```

```
    sample(assign) #randomise order of assignment within each
stratum

  }) %>% ungroup()


table(data$strata,data$treat) #check allocation


## Creating the risk list in each treatment group and randomly
assigning intervention which is 50/50


prop4<- 0.5

prop5<-0.5


#At Risk in Treatment Group 1 (WHO5 GROUP)


tg1<- subset(data,data$treat==1 & data$WHO5<50)


tg1$strata<-ifelse(tg1$WHO5>=median(tg1$WHO5),1,2)


tg1<-tg1 %>% group_by(strata) %>% mutate(nudgetreat={

  n <- n()

  allocation1 <- floor(n*prop4) #proportion of stratum
allocated to arm1

  allocation2 <- floor(n*prop5) #proportion of stratum
allocated to arm2

  assign <- c(rep(1, allocation1),  rep(0,n- (allocation1)))
#vector of 1s (treat) and 0s (control)

  sample(assign) #randomise order of assignment within each
stratum
```

```
}) %>% ungroup()



#Treatment Group 2 (LA GROUP)


tg2<-subset(data,data$treat==2& data$LA_OCT<0.75)

tg2$strata<-ifelse(tg2$LA_OCT>=median(tg2$LA_OCT),1,2)


tg2<-tg2 %>% group_by(strata) %>% mutate(nudgetreat={

  n <- n()

  allocation1 <- floor(n*prop4) #proportion of stratum
allocated to arm1

  allocation2 <- floor(n*prop5) #proportion of stratum
allocated to arm2

  assign <- c(rep(1, allocation1),  rep(0,n- (allocation1)))
#vector of 1s (treat) and 0s (control)

  sample(assign) #randomise order of assignment within each
stratum

}) %>% ungroup()



riskdata<-rbind(tg1,tg2)


#create nudgelist by merging riskdata with those not at risk
in TG1 or TG2 but who are receiving an email


noriskdata<-subset(data, data$treat !=0)
```

```
noriskdata<-noriskdata %>%
  anti_join(riskdata,by = "Rand_code")
noriskdata$nudgetreat<-2


nudgelist <-rbind(noriskdata,riskdata)


nudgelist <- nudgelist %>%
  mutate(treat = case_when(
    treat == 1 ~ "WHO5",
    treat == 2 ~ "LA",
    TRUE ~ as.character(treat)  # Keep other values as they
are
  ))


nudgelist <- nudgelist %>%
  mutate(nudgetreat = case_when(
    nudgetreat == 0 ~ " 1:1 ",
    nudgetreat == 1 ~ "SilverCloud",
    nudgetreat == 2 ~ "KBA",
    TRUE ~ as.character(nudgetreat)  # Keep other values as
they are
  ))



write.csv(nudgelist,"nudgelist.csv", row.names = FALSE)
```

```
table(riskdata$strata,riskdata$treat) #check allocation


table(nudgelist$nudgetreat,nudgelist$treat) #check allocation


#when this code was run at 13:08 on 16/10/2024 the
intervention list looked as follows:-

#            LA TG   WHO5 TG

#1:1          205     336

#KBA emails   3934    3672

#SilverCloud  204     335
```