Evaluation of an Integrated Mental Health Liaison Service (RAID)

Prepared for the Northern Health and Social Care Trust

April 2018

The Strategy Unit.

NHS
Midlands and Lancashire
Commissioning Support Unit

About this report

This report looks at the effect of an enhanced 24/7 mental health liaison service. The new service—called RAID (Rapid Assessment Interface Discharge)—was launched in June 2015 at two acute hospital sites run by the Northern Health and Social Care Trust, one of five health and social care trusts in Northern Ireland.

The report outlines the key elements of the RAID service, then describes the study design and methods used to undertake a quantitative analysis comparing outcomes for patients treated by the RAID service with a matched control group. The report describes the results of the analysis and discusses the findings.

The report was produced by the Strategy Unit, a specialist team providing research, analysis and strategic change expertise from within the NHS. If you are interested in finding out more about the project or would like to discuss similar work please contact Paul Seamer, Analytics Manager, the Strategy Unit.

Paul Seamer | Analytics Manager
Strategy Unit
paulseamer@nhs.net

Contents

Key findings

A high proportion of patients treated for physical health conditions also have co-morbid mental health problems, and there is increasing acceptance of a need to improve the awareness and treatment of mental health problems within acute hospitals.

The Northern HSC Trust RAID service is a specialised multidisciplinary liaison psychiatry and psychological medicine team operating in two acute hospitals (Antrim Area and Causeway hospitals). The RAID model involves rapid assessment of patients presenting in ED and those admitted to general hospital wards so that they can receive appropriate interventions for their physical and mental health, either in the community or in hospital.

We used a retrospective matched cohort study design to compare likelihood of admission from the emergency department and average length of stay for patients treated by RAID with a matched comparison group of similar patients treated in other acute hospitals in the region. For patients treated by the RAID service in the emergency department their average risk (or chance) of admission was 18.7% lower (95% confidence interval –9.3% to –27.3%). Average length of stay for a wider group of patients (including those benefitting indirectly from RAID) treated on general wards was 10.8% lower (95% confidence interval –9.2% to –12.4%). Both findings were statistically highly significant ($P < 0.001$).

Key findings

Caution is required before inferring causality from non-experimental study designs. Although the matched comparison group had similar age, gender, arrival mode or method of admission, diagnoses, prior hospital utilisation to the comparison group of patients they may have differed in unobserved ways that could have influenced outcomes.

The small number of previous studies that investigated the effectiveness of the RAID model suggested the potential for the model to deliver both improved outcomes and cost savings. Our findings are similarly positive and indicate that the RAID service in the Northern HSC Trust has led to better outcomes and lower resource use in two acute hospitals.

Study background

About the Northern Trust

The Northern Health and Social Care Trust (NHSCT) is one of five Health and Social Care (HSC) Trusts in Northern Ireland. HSC Trusts are the main providers of health and social care in Northern Ireland. The NHSCT was established following the merger of three former Trusts—Causeway, Homefirst and United. The NHSCT provides a comprehensive range of health and social care services to a population of almost 436,000 people across a geographical area spanning four new council areas—Antrim and Newtownabbey Borough Council, Causeway Coast and Glens Borough Council, Mid and East Antrim Borough Council, and Mid Ulster District Council.

Emergency hospital services are provided on two sites, Antrim Area Hospital and Causeway Hospital in Coleraine. Both hospitals provide a range of services including a consultant-led 24-hour emergency care department.

About the RAID service

The NHSCT RAID service is a specialised multidisciplinary liaison psychiatry and psychological medicine team operating in two acute hospitals (Antrim Area and Causeway hospitals). The service is modelled on and shares a high degree of fidelity with the original RAID service model developed in Birmingham, England.

The RAID model involves rapid assessment of patients presenting in ED and those admitted to general hospital wards so that they can receive appropriate interventions for their physical and mental health, either in the community or in hospital. The service operates 24 hours a day, seven days a week and has a target response time for assessments in the emergency department of two hours, and 24 hours for ward assessments. It is available to any adult patient in the hospital who presents with self-harm, has problems with dependence or misuse of alcohol or other substances, has mental health problems associated with old age or has mental illness co-existing with physical illness. There are no rigid referral criteria and hospital staff are encouraged to refer any patients who may benefit from the service.

The RAID team provide close clinical support and supervision in mental health interventions for general hospital staff and actively seek to raise awareness of mental health issues within the hospital. The team also deliver formal and informal training to hospital staff across a range of mental health issues.

Previous research on RAID

There is an accumulating body of evidence that shows a high proportion of patients with physical health conditions also have co-morbid mental health problems. The government's 'No Health Without Mental Health' strategy highlighted the need to improve the awareness and treatment of mental health issues in acute hospitals (1). In response to these challenges, a new service—the Rapid Assessment, Interface and Discharge Service—or RAID—was developed to provide greater access to mental health expertise for patients and staff in acute hospitals. The first RAID service was established in City Hospital in Birmingham (England) in 2009, and was subsequently rolled out to other sites across the West Midlands.

Parsonage and Fossey identified the following key features of the initial RAID service (2):

- Provision of a comprehensive range of mental health specialties within one multi-disciplinary team
- Able to provide a rapid response 24 hours a day, 7 days a week
- Able to meet the mental health needs of all adult patients in the hospital
- Provision of formal teaching and informal training to staff throughout the hospital
- An emphasis on diversion and discharge from the emergency department and on facilitating early but effective discharge from general wards
- Provision of follow-up clinics for patients discharged from hospital

(1) Department of Health. *No health without mental health: a Cross-Government Mental Health Outcomes Strategy for People of All Ages*. Available from: https://www.gov.uk/government/publications/no-health-without-mental-health-a-cross-government-mental-health-outcomes-strategy-for-people-of-all-ages-a-call-to-action [accessed 1st April 2018].

(2) Parsonage M, Fossey M. *Economic evaluation of a liaison psychiatry service*. London: Centre for Mental Health; 2011.

Previous research on RAID

A small number of studies have investigated the effectiveness and cost-effectiveness of RAID services. The findings of these studies suggest the potential for the RAID model to deliver both improved outcomes and cost savings.

An economic evaluation of the pilot service at City Hospital in Birmingham found significant economic savings—a benefit-cost ratio of 4:1—resulting from the success of the service in promoting quicker discharge from hospital and fewer re-admissions (1). The evaluation was based on critical scrutiny and re-analysis of data collected as part of a wider internal review, which later formed the basis for a separate study (2).

An evaluation of the subsequent roll-out of the RAID service to other acute hospitals in the West Midlands region found patients seen by RAID in the emergency department were significantly less likely to be admitted than a group of matched control patients but spent significantly more time in the department than controls; and were more likely to re-attend within 28 and 90 days. Patients seen by RAID on inpatient wards were found to have a significantly lower average length of stay than controls (3).

(1) Parsonage M, Fossey M. *Economic evaluation of a liaison psychiatry service*. London: Centre for Mental Health; 2011.

(2) Tadros G, et. al. Impact of an integrated rapid response psychiatric liaison team on quality improvement and cost savings: the Birmingham RAID model. *The Psychiatrist*. 2013;37(1): 4–10.

(3) The Strategy Unit. *Rapid Assessment Interface And Discharge Liaison: Economic Evaluation of the Birmingham and Solihull Roll-Out*. Birmingham: The Strategy Unit; 2012.

Study design

Approach

Random assignment is considered the gold standard in empirical evaluation work, but for many reasons it is not always practical or feasible to implement a random assignment study. In a random assignment study to determine the effect of RAID, subjects would be randomly assigned to a treatment group that would receive the RAID service and a control group that would not receive RAID. The process of randomisation ensures any difference in outcomes between the two groups can be interpreted as a 'RAID effect'. In some circumstances, where an experimental design is not feasible, non-experimental methods can provide causal estimates that are comparable to those obtained from experiments.

This analysis was a retrospective matched cohort study comparing likelihood of admission from the emergency department and average length of stay for patients treated by RAID with a matched control group.

Cases—patients treated by RAID during the study period at the two intervention sites [1]

Controls—patients similar to cases at baseline, but who presented at other acute hospitals in the region during the study period [2]

[1] The cohort of patients seen by RAID on general wards was expanded to include all patients with a mental health diagnosis irrespective of whether they received treatment from the RAID team. This was done because the RAID team provided training and support to hospital staff managing these patients during the study period.

[2] To the best of our knowledge none of the other acute hospitals in the region made any changes to their liaison mental health services during the study period. We assume that 'usual care' at the non-intervention sites may have involved some liaison mental health, but at a reduced level, broadly equivalent to that available in the Northern Trust prior to the introduction of RAID.

Approach

Cases and controls were ‘matched’ on a set of characteristics thought to influence or predict outcomes from treatment. In this way, the outcomes for controls can be thought of as a proxy for the outcomes of cases if they had not received the intervention. The outcomes of cases and controls were compared with a view to attributing any differences to a ‘RAID effect’.

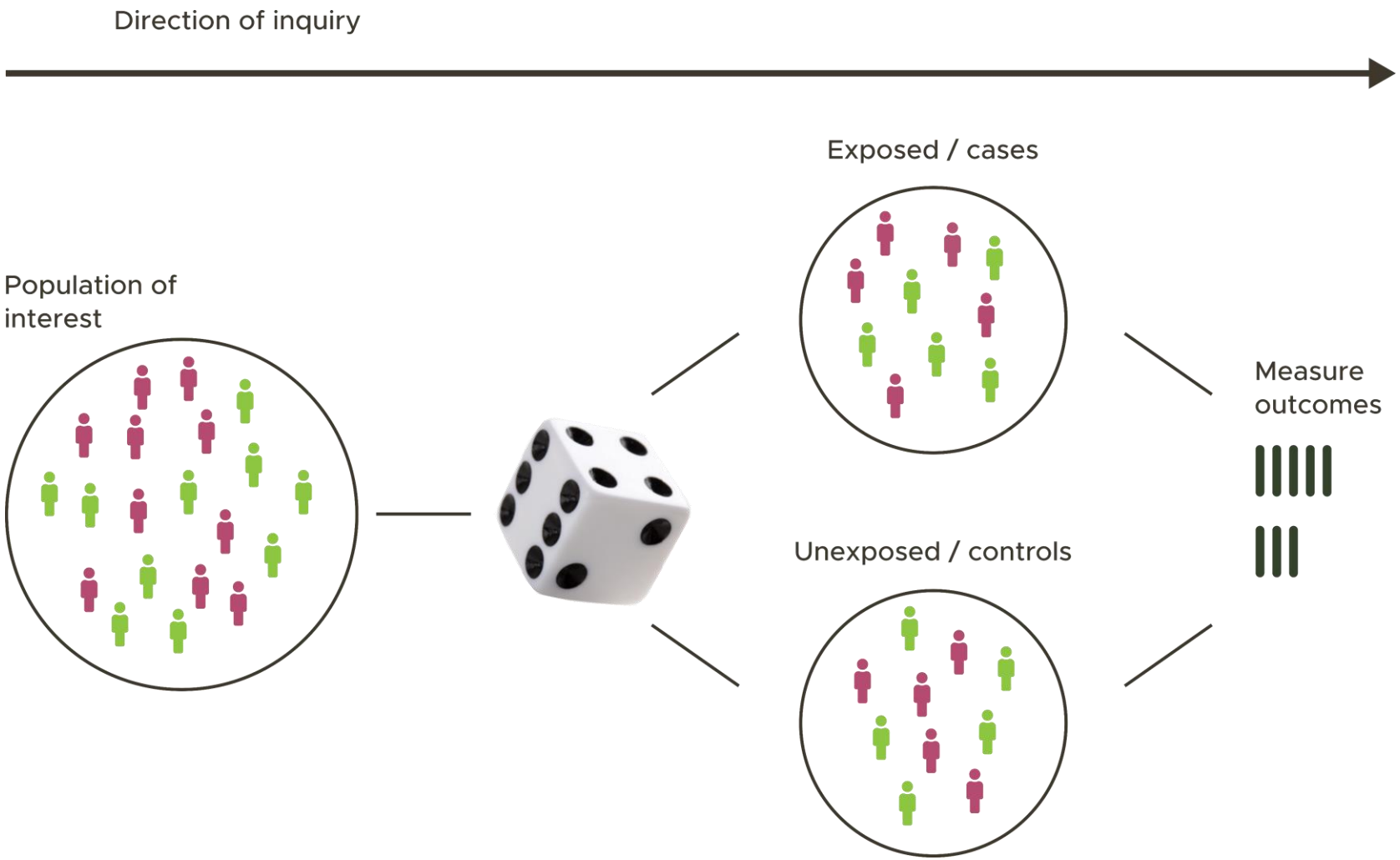
Main outcomes

The primary outcomes for this study were:

- Likelihood of admission for patients presenting in ED
- Average length of stay for patients admitted to general hospital wards

The avoidance of unnecessary admissions and the facilitation of early but effective discharge are important aims for the RAID service and can also be measured with acceptable accuracy from routine datasets. There are many other potential measures / outcomes that might plausibly be changed through improved liaison mental health services but these are beyond the scope of this evaluation.

The Gold Standard: a randomized experiment

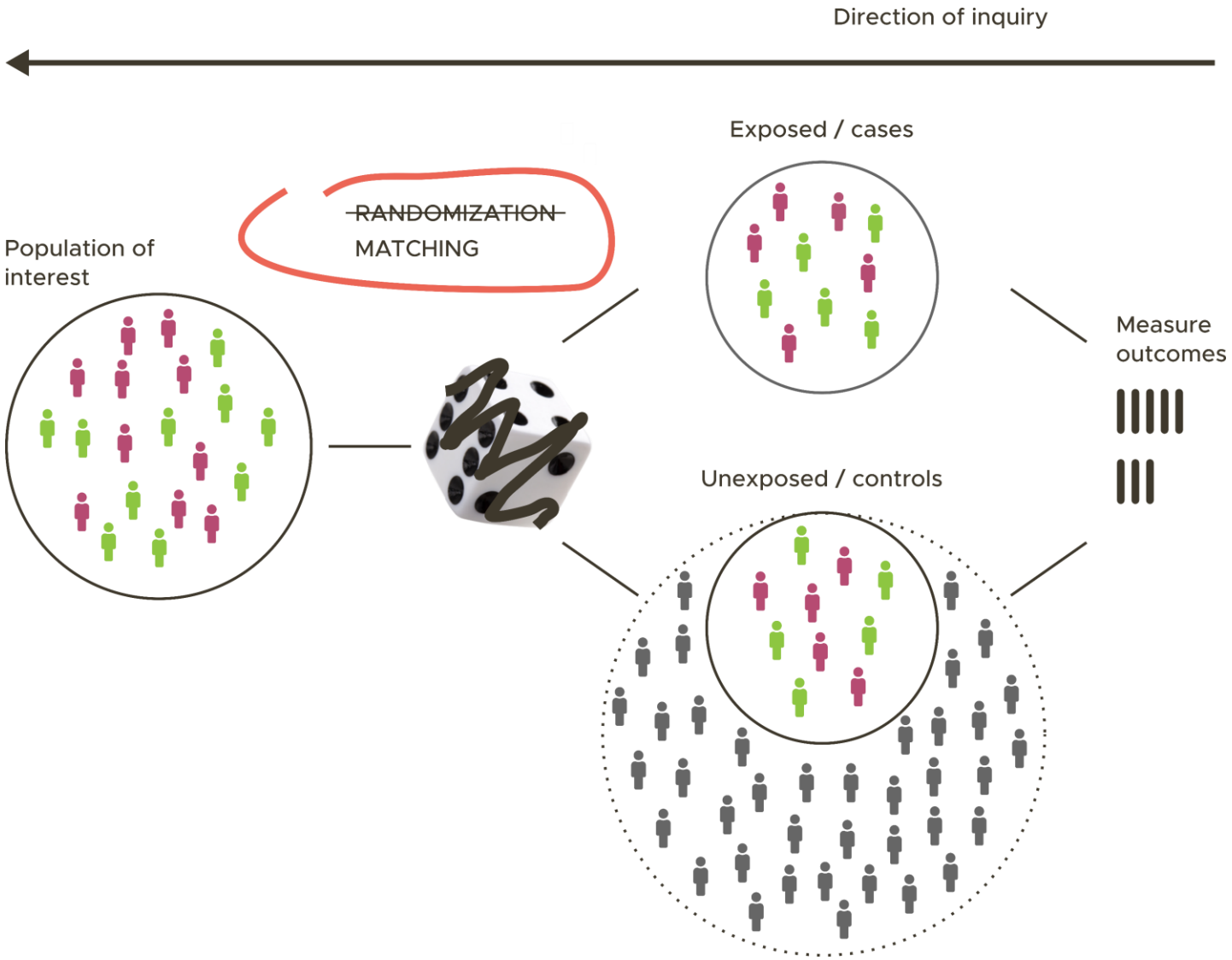


There is widespread acceptance that for inferring cause and effect randomized experiments are preferable to other methods. However, for a variety of ethical and practical considerations an experiment is not always feasible.

Important features of a randomized experiment

- Objectivity of treatment assignment (unbiased estimates)
- Balance on covariates (observed and unobserved)
- Prospective—no access to outcome data (prevents cheating / p-hacking)

The retrospective matched cohort design (a quasi-experiment)



In situations where an experiment is not feasible it is sometimes possible to design a comparative observational study that approximates a hypothetical randomized experiment. If carefully conducted such studies can provide reliable inferences on cause and effect.

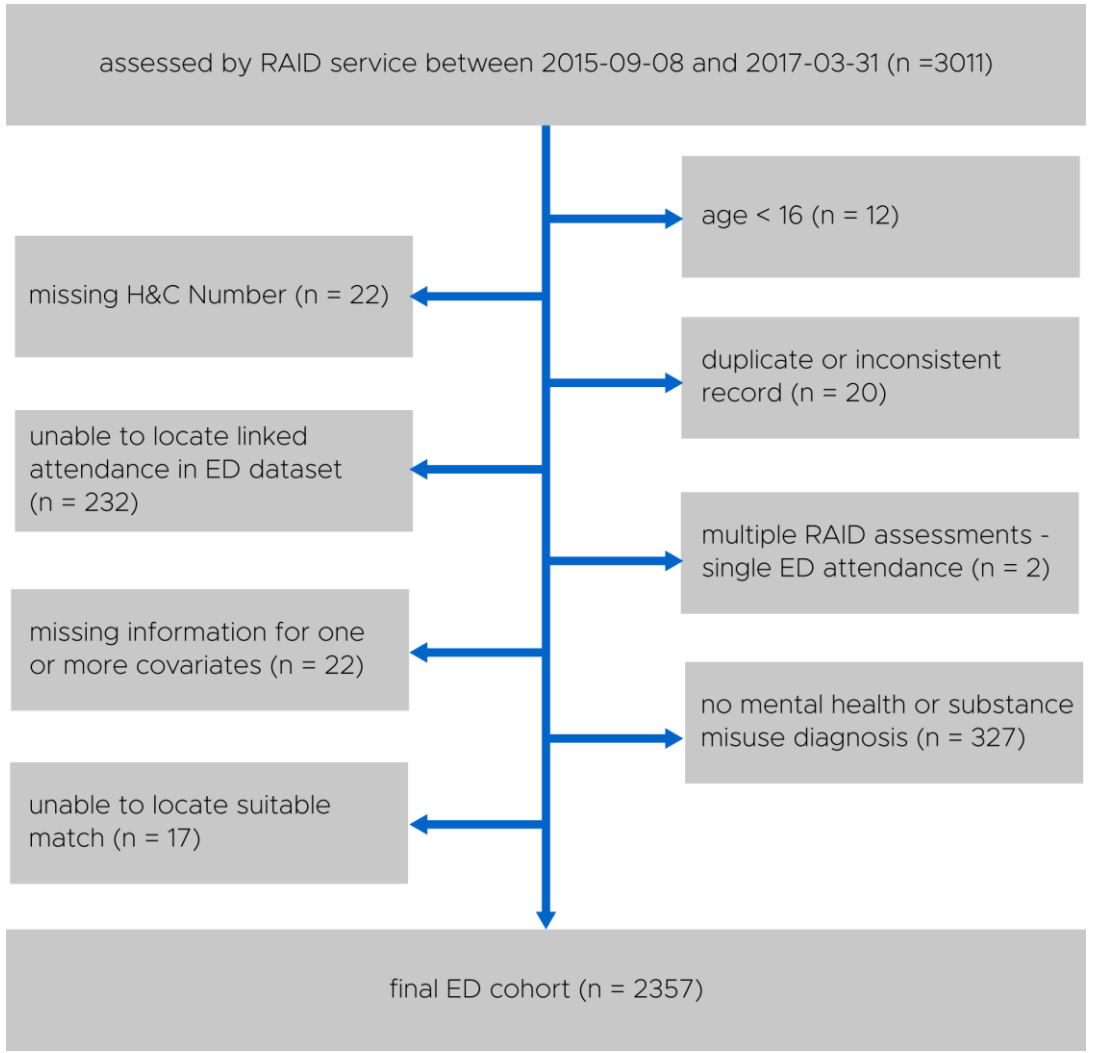
Important questions to ask about observational study designs

- What was the hypothetical randomized experiment that led to the observed dataset
- Are sample sizes in the dataset adequate
- Who are the decision makers for treatment assignment and what measurements where available to them?
- Are the key covariates measured well?
- Can balance be achieved on key covariates?

Study population

Study population—ED cohort

We identified all patients referred to the RAID team from the emergency department with a completed assessment during the period 8 September 2015 to 31 March 2017. Some patients referred to the RAID service refuse an assessment or leave the department before an assessment can be completed—these patients were excluded from the analysis as our interest was estimating the average treatment effect on the treated. This parameter represents the average effect of the intervention among those who have been exposed to it.

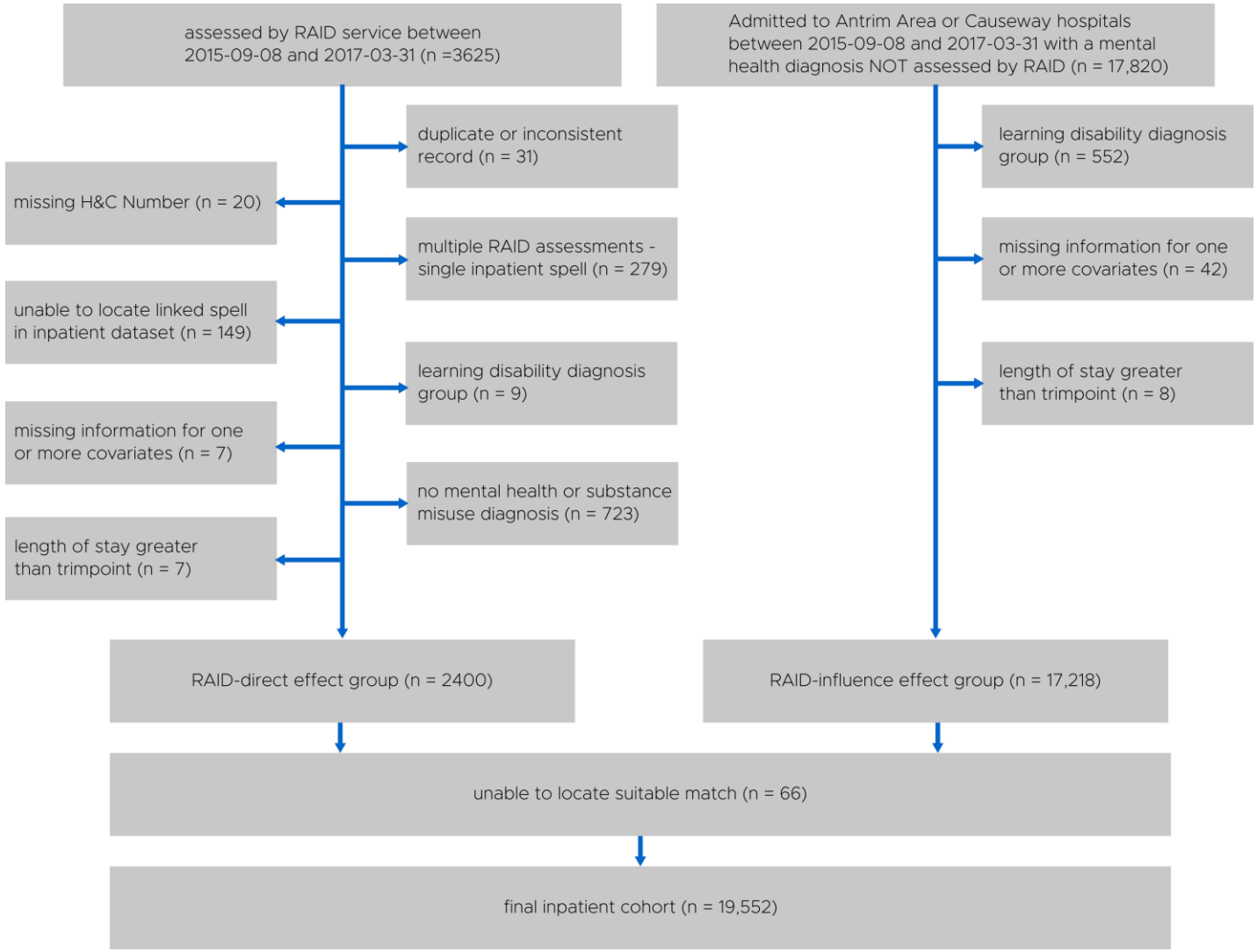


Study population—inpatient cohort

We identified all patients referred to the RAID team from a general ward with a completed assessment during the period 8 September 2015 to 31 March 2017. We refer to this group as the RAID-direct group.

We also identified a second group of patients who were admitted with a mental health or substance misuse diagnosis to one of the intervention hospitals but were not assessed by RAID. We refer to this second group as the RAID-influence group.

Patients referred to the RAID service from the observation ward at Antrim Area Hospital were excluded (Causeway Hospital does not have an observation ward).



Data sources & linkage

Data sources

Three main datasets were used in this evaluation:

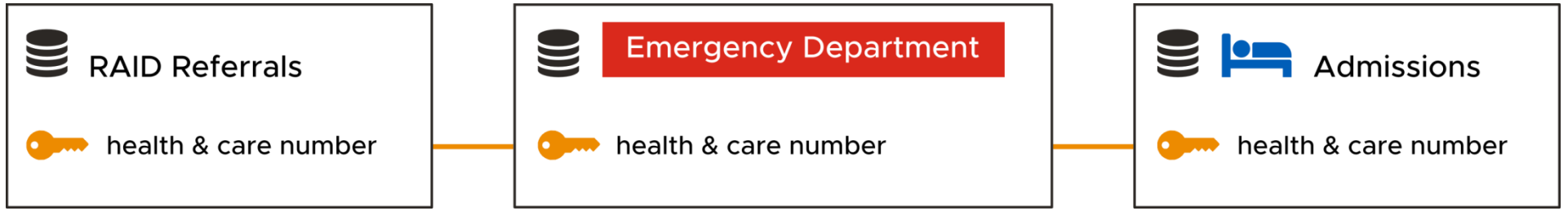
1. RAID Referrals
2. Emergency Department
3. Admissions

Patient level information on referrals to the RAID service was supplied by the Northern Trust. The regional emergency department and admissions datasets were sourced from the Honest Broker Service (HBS), part of the Regional Business Services organisation. The HBS manages a regional data warehouse that securely stores patient-level data submitted from each health and social care trust. A key aim of the HBS is to facilitate the sharing of non-identifiable data for health and social care related research.

Data linkage

A process of pseudonymisation was used to replace the health and care number of patients in the three main datasets with a unique identifier (key) which does not reveal their 'real world' identity but allows researchers to track individuals over time and across multiple datasets.

The pseudonymisation was performed by the HBS and the pseudonymised datasets, with all personal or identifiable data removed, were supplied to the Northern Trust. The HBS terms of establishment allow for the sharing of data between the HBS and all regional health and social care organisations. The three datasets were subsequently shared with the Strategy Unit by the Northern Trust under a separate Data Access Agreement (DAA) permitting their time limited use for the purpose of this project.



Selecting a comparison group

Selecting a comparison group

For efficient causal inference it is important to compare treated and control groups that are as similar as possible. For the RAID comparison group we sought to identify patients from other acute hospital sites in the region that would have been eligible for the RAID service had such a service been available.

An important element of the RAID service is the absence of any rigid referral criteria—hospital staff are encouraged to refer anyone who they think may have mental health problems or substance misuse issues. We sought to identify all patients where a potential mental health problem or substance misuse issue was indicated in their hospital record.

The emergency care departments at Antrim Area and Causeway hospitals both provide a consultant-led 24-hour service. To ensure comparability we selected as potential control sites only hospitals with designated type-1 emergency care departments. Belfast City hospital was added to the set of control sites as it is the largest hospital in the region and until 2011 operated a type-1 emergency care department. This meant a maximum of eight control sites to provide a 'pool' of comparison patients from which the matched comparisons could be selected.

For the ED cohort, it was necessary to further restrict the number of control sites. Two different emergency department software systems are used in the region; major differences in the way information is recorded by the two systems meant we were only able to consider control sites using the same system as the intervention hospitals.

Intervention and control hospitals

	ED cohort	Inpatient cohort
Belfast Trust		
Belfast City	.	✓
Mater	✓	✓
Royal Victoria	✓	✓
Northern Trust		
Antrim Area	RAID	RAID
Causeway	RAID	RAID
South Eastern Trust		
Ulster	.	✓
Southern Trust		
Craigavon Area	.	✓
Daisy Hill	.	✓
Western Trust		
Altnagelvin Area	✓	✓
South West Acute	✓	✓

Selecting a comparison group

Not all patients treated by the RAID team have a mental health problem or substance misuse issue indicated on their hospital record; and only a proportion of all patients with a mental health problem or substance misuse issue indicated on their hospital record are referred to the RAID service.

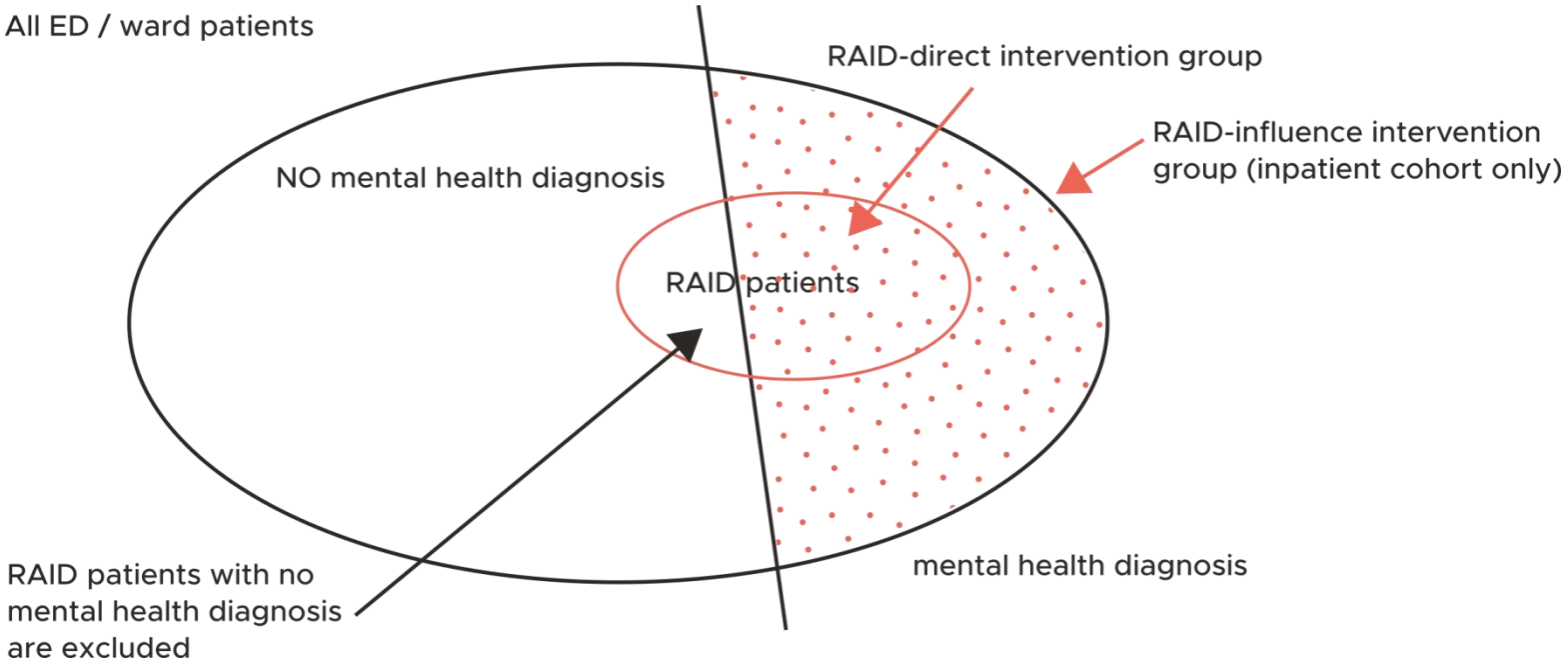
This raises two important questions:

1. the extent to which patients with a mental health diagnosis and treated by RAID differ from those with a mental health diagnosis but not referred to RAID; and
2. the extent to which patients without a mental health diagnosis and treated by RAID differ from those without a mental health diagnosis but not referred to RAID.

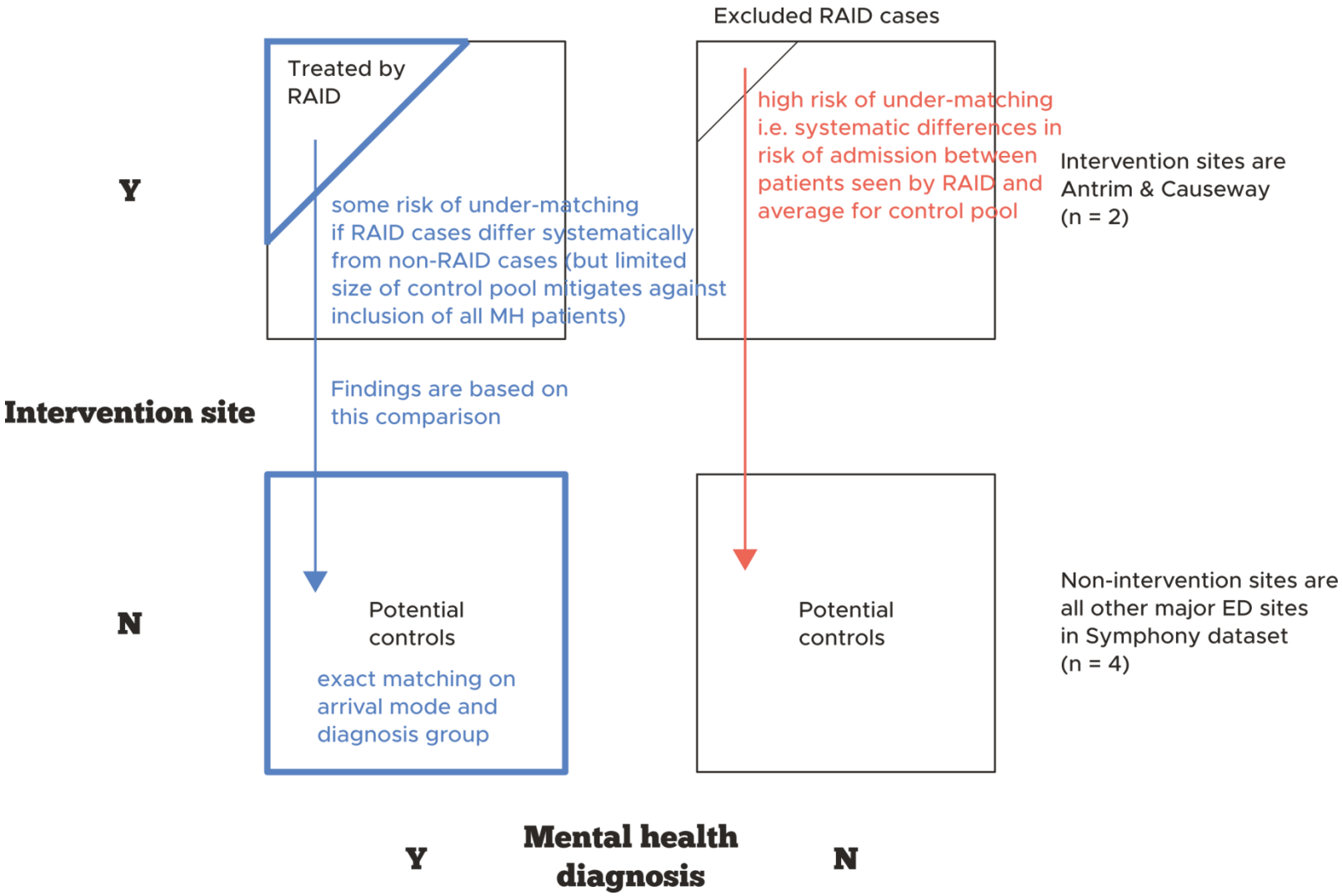
Of particular concern for this study is the risk of systematic under-matching—a failure to select controls that are sufficiently like the cases. It is easy to hypothesise that in both cases the patients seen by the RAID team are systematically different to the wider group of patients not seen by RAID. For example, it is likely that the prevalence of low level mental health or substance misuse issues is higher among patients referred to RAID but without a documented diagnosis of mental illness than in the wider hospital population of patients without a diagnosis of mental illness. For this reason, patients seen by RAID without a documented diagnosis of mental illness were excluded.

Selecting a comparison group

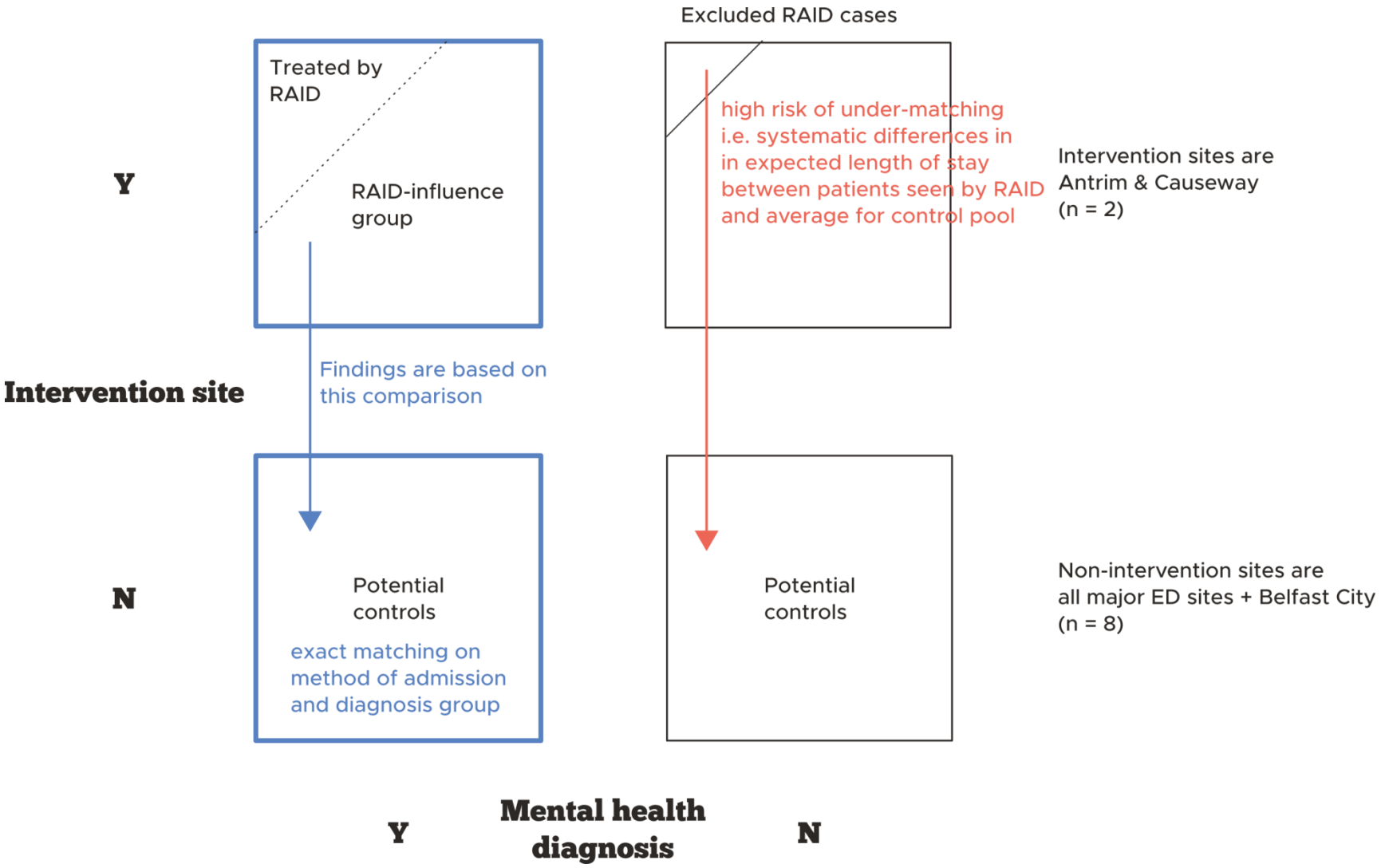
For the inpatient cohort, the inclusion of the RAID-influence group has the added benefit of mitigating the risk of under-matching—the comparison made is for all patients with a mental health diagnosis not just those treated by RAID. For the ED cohort, we judged the risk of under-matching to be reduced because decision-making is more immediate (more random) in that environment. But in any case, the smaller number of control sites would have limited the closeness of matches achievable if the group had been widened to include patients not seen by RAID.



ED cohort—comparison group



Inpatient cohort—comparison group



Identifying patients with a mental health diagnosis

We planned to use information extracted from the hospital patient administration system (PAS) and emergency department software system to identify patients with a mental health diagnosis—as a proxy for RAID eligibility. Clinical coding staff use information from patient notes to code details of a patient's diagnoses using the ICD-10 classification. Information is recorded for a patient's primary diagnosis (the most serious or resource intensive diagnosis during the hospitalisation) and if present multiple secondary diagnoses. Chapter 5 of ICD-10 covers diagnoses of mental and behavioural disorders. We used the presence of a mental or behavioural disorder diagnosis (primary or secondary, and in any episode of a continuous inpatient spell) to identify patients for the RAID influence group, and for our pool of potential comparison patients. These patients were then sub-divided into one of ten diagnostic groups (see slide 31).

In the emergency department setting there tends to be greater variability in the extent of information collected and less standardisation in the way information is recorded. Some trusts in the region, including the two intervention sites use the Symphony software system, but others use the Northern Ireland Regional Accident and Emergency System (NIRAES). The extracts we obtained from the NIRAES system did not contain any information on diagnoses meaning we were unable to identify patients who may have been referred to RAID. For this reason, for the ED cohort we excluded from the set of control sites hospitals using the NIRAES system.

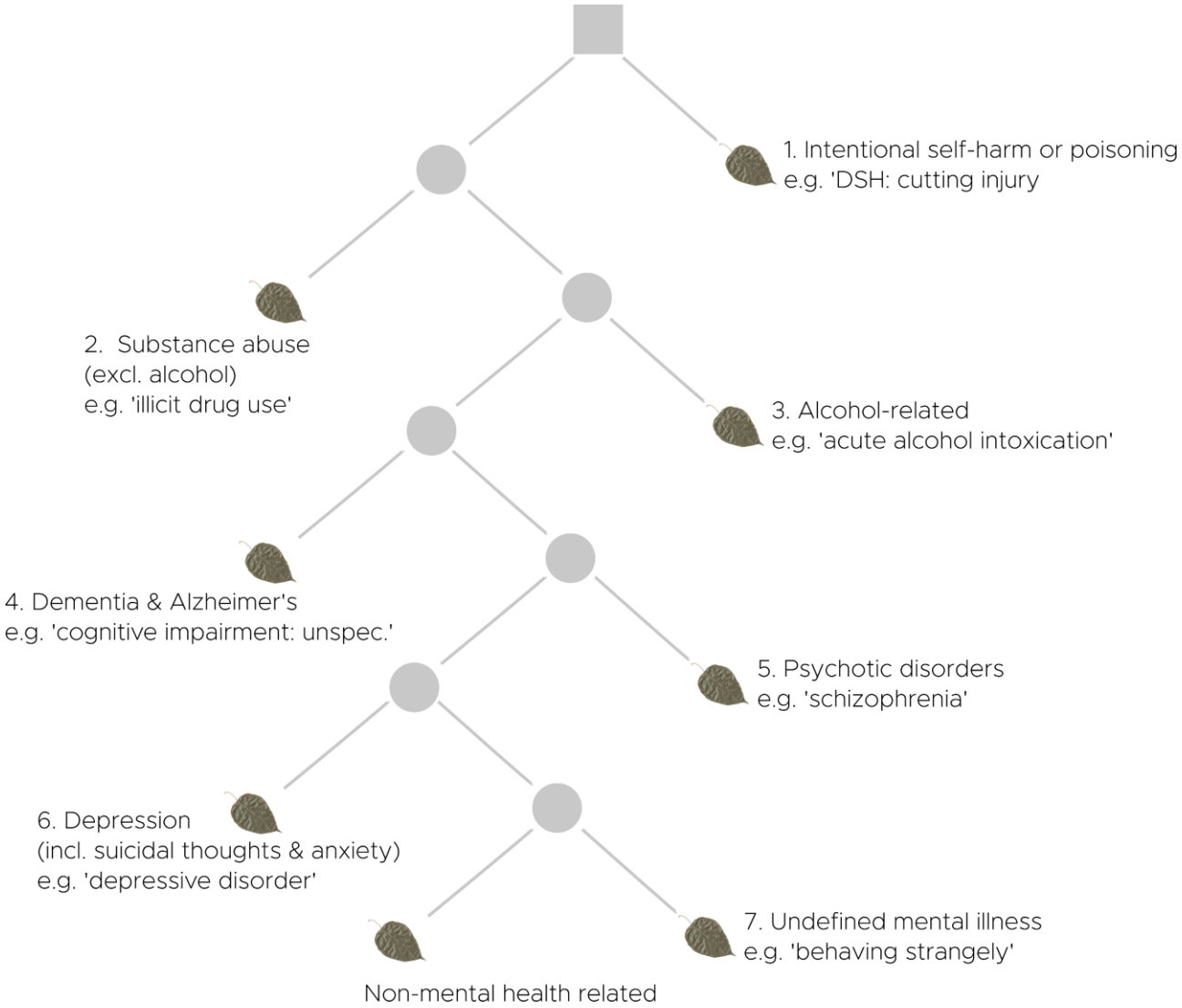
Information on diagnosis for patients presenting in ED

In the extracts obtained from the Symphony system information about a patient's diagnosis or presenting complaint was stored as unstructured text. The high degree of variability in the words and phrases used to record a diagnosis or complaint and the size of the dataset meant that manual coding of individual responses was impractical. Text mining methods, however, can be used to process large amounts of unstructured textual information and summarise or classify the contents. We used text mining techniques to locate keywords or phrases relating to the most common reasons for referral to the RAID team. Based on the presence of particular keywords or phrases in the diagnosis and complaint fields we allocated patients to one of seven broad diagnostic groups (see following slide).

Example information from Symphony ED dataset

pseudo HCN	gender	age	diagnosis / triage complaint
.	M	28	Acute alcohol intoxication
.	F	19	History of depression and under care of YPC. Self harm wounds to left wrist 1/7 ago . . . . Gp referral for further assessment of wounds. States happy with care provided by YPC
.	M	42	States took librium ?8tabs +diazepam 5mgs x? 6, also 15 pints larger onboard. Pt low mood wants to kill himself
.	M	35	Anxiety State

Decision tree for assignment to ED diagnostic groups

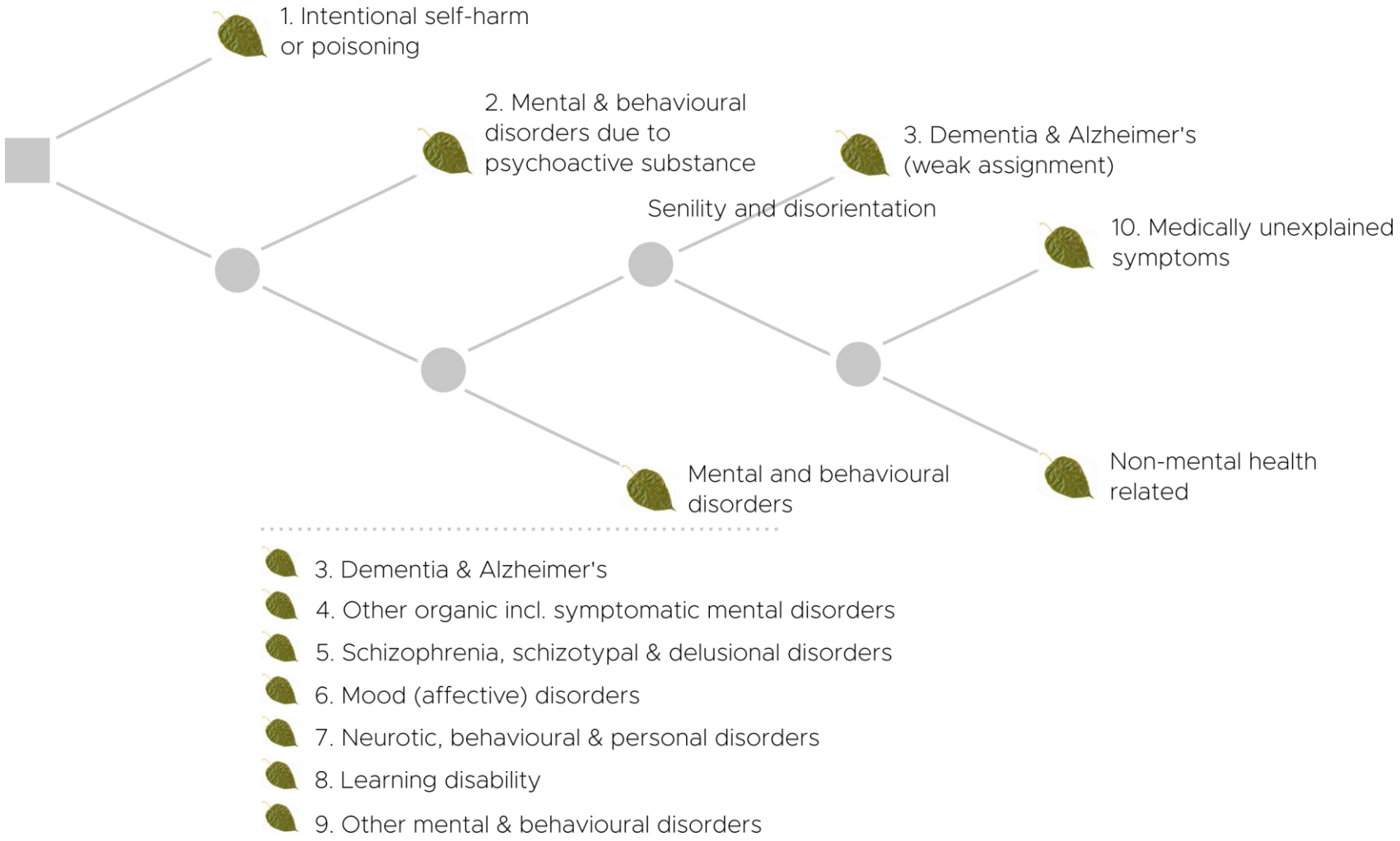


Algorithm for assignment to a diagnosis group

A set of rules was developed to allocate patients (RAID cases and potential controls) to a diagnostic group. Assignment was determined by the presence of keywords or phrases in the textual information recorded about a patient's diagnosis or complaint. The following hierarchy (which was determined by evaluating the implications of the diagnosis / complaint and the relevance and specificity of the words / phrases) was used. Assignment was on a 1:1 basis i.e. a patient could only be allocated to a single group.

1. Intentional self-harm / poisonings
2. Substance abuse (excl. alcohol)
3. Alcohol-related
4. Dementia & Alzheimer's
5. Psychotic disorders
6. Depression (incl. suicidal thoughts & anxiety)
7. Undefined mental illness

Decision tree for assignment to inpatient diagnostic groups



Algorithm for assignment to a diagnosis group

A set of rules was developed to allocate patients (RAID cases and potential controls) to a diagnostic group. Assignment was determined by the presence of a relevant ICD-10 diagnosis code in any episode in the spell. The following hierarchy (which was determined by evaluating the implications of the diagnosis and the relevance and specificity of the codes) was used. Assignment was on a 1:1 basis i.e. a patient could only be allocated to a single group.

1. Intentional self-harm / poisonings (any code in any position)
2. Mental and behavioural disorders due to psychoactive substance group (any code in any position)
3. Mental & behavioural disorders (if more than one relevant code, allocated to a group based on position within the diagnosis codes)
4. Senility and disorientation (any code in any position)
5. Medically unexplained symptoms (any code in any position)

Matching methods

Matching methods

A key benefit of randomization is that the treated and control groups are guaranteed to be only randomly different from one another. Matching attempts to replicate a randomized experiment as closely as possible by obtaining treated and control groups with similar distributions (or 'balance') of all variables that are possibly predictive of the outcome under study. These variables are termed 'covariates'.

Variable selection was performed blind of the observed outcomes. Choice of matching variables was based on previous research and availability of information from the datasets used. The following slides list the matching variables for used for the two cohorts.

Matching variables—ED cohort

Patient characteristics

- gender
- age
- deprivation

Attendance characteristics

- incident type (self-harm, RTA, etc.)
- attendance type (unplanned, planned)
- arrival mode [1]

Presentation characteristics

- triage category
- diagnosis group [1]

Prior utilisation

All prior utilisation measures were calculated separately for both previous 28 days (4 weeks) and previous 180 days (6 months). Twelve measures in total.

- count of ED attendances
- count of ED attendances ending in admission
- count of unplanned admissions
- count of planned admissions
- sum of unplanned hospital bed-days
- sum of planned hospital bed-days

Environmental factors

- time of day (day/night)
- day of week
- month of year

[1] Exact matching used.

Matching variables—inpatient cohort

Patient characteristics

- gender
- age
- deprivation

Admission characteristics

- method of admission [1]

Presentation characteristics

- number of comorbidities
- number of distinct treatment specialties in spell
- diagnosis group [1]

Prior utilisation

All prior utilisation measures were calculated separately for both previous 28 days (4 weeks) and previous 180 days (6 months). Twelve measures in total.

- count of ED attendances
- count of ED attendances ending in admission
- count of unplanned admissions
- count of planned admissions
- sum of unplanned hospital bed-days
- sum of planned hospital bed-days

Environmental factors

- time of day (day/night)
- day of week
- month of year

[1] Exact matching used.

Finding matches

From the wider set of patients in the control hospitals, we selected a matched subset that was similar to the RAID cases with respect to variables that were likely to be predictive of outcomes or assignment to treatment group. For some variables we required that the corresponding RAID and matched comparison patients were exactly equal.

Cohort	Variables	Method of matching
ED	Arrival mode, and diagnosis group	Exact
Inpatient	Method of admission, and diagnosis group	Exact

Matched comparison patients were selected using genetic matching (1). Genetic matching is a multivariate matching method that uses an evolutionary [1] search algorithm to determine weights for a distance measure (on covariates) between treated and untreated cases that optimises post-matching balance. Simulation studies have shown genetic matching outperforms alternative methods across a range of possible covariate distributions (2). We selected one matched comparison patient for each RAID patient, since this results in better balance than one-to-many (or ratio) matching. Matching was done with replacement, meaning that the same comparison patient might be matched to more than one RAID patient. Matching with replacement ensures each treated case is matched to the most similar untreated case and therefore produces greater bias reduction than matching without replacement.

[1] An evolutionary algorithm (EA) uses a collection of heuristic rules to modify a population of trial solutions in such a way that each generation of trial values tends to be, on average, better than its predecessor.

(1) Sekhon JS. Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R. *Journal of Statistical Software*. 2011;42(7): 1–52.

(2) Diamond A, Sekhon JS. Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. *Review of Economics and Statistics*. 2013;95(3): 932–945.

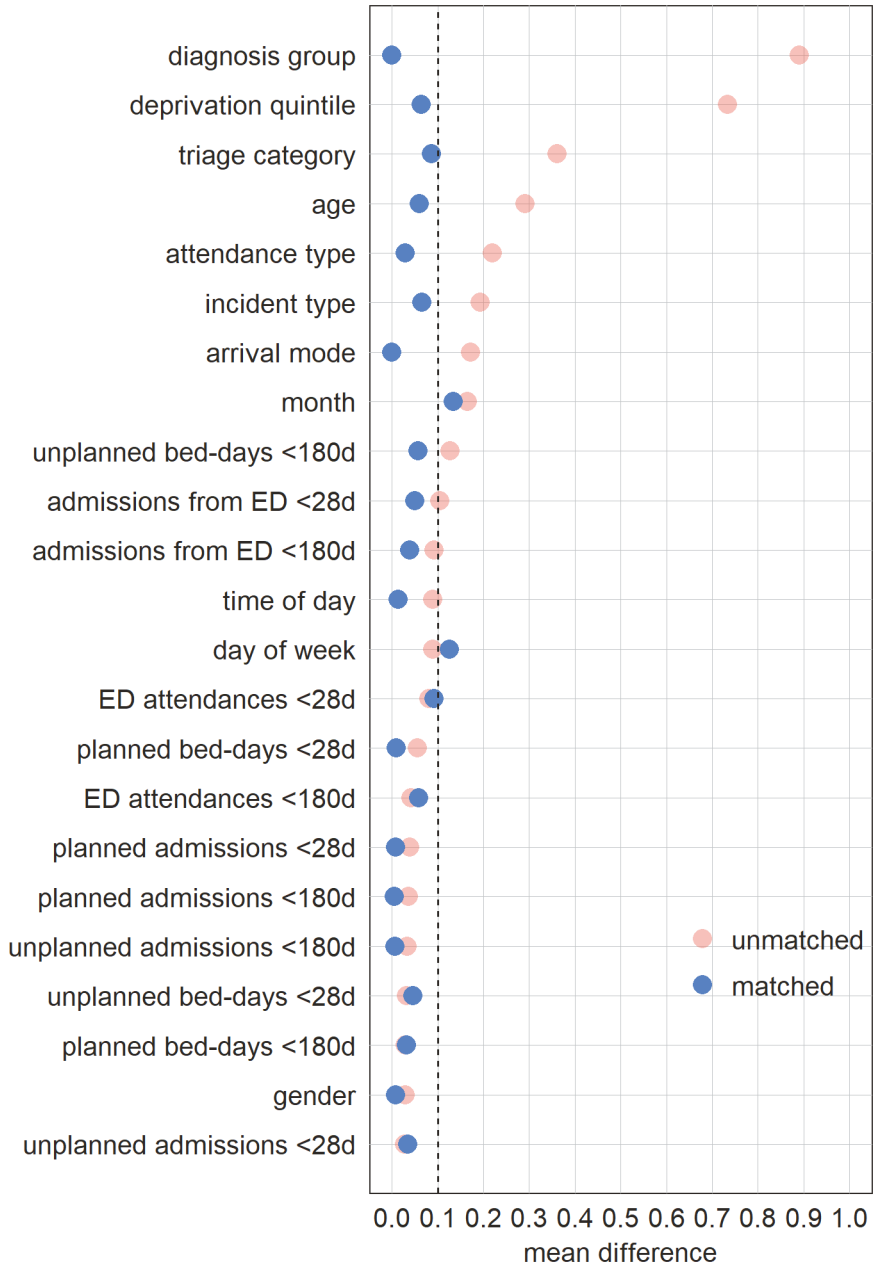
Estimating the effect size of RAID

Comparisons between RAID cases and the matched comparison group were made using multivariable regression. The purpose of the regression models was to adjust for any small differences in baseline characteristics still present after matching. The regression models produced a ‘best estimate’ of the relative difference in the chance of admission between RAID cases seen in ED and the matched comparison group, and in length of stay between RAID cases seen on general wards and the matched comparison group.

Results

Assessment of matching—ED cohort

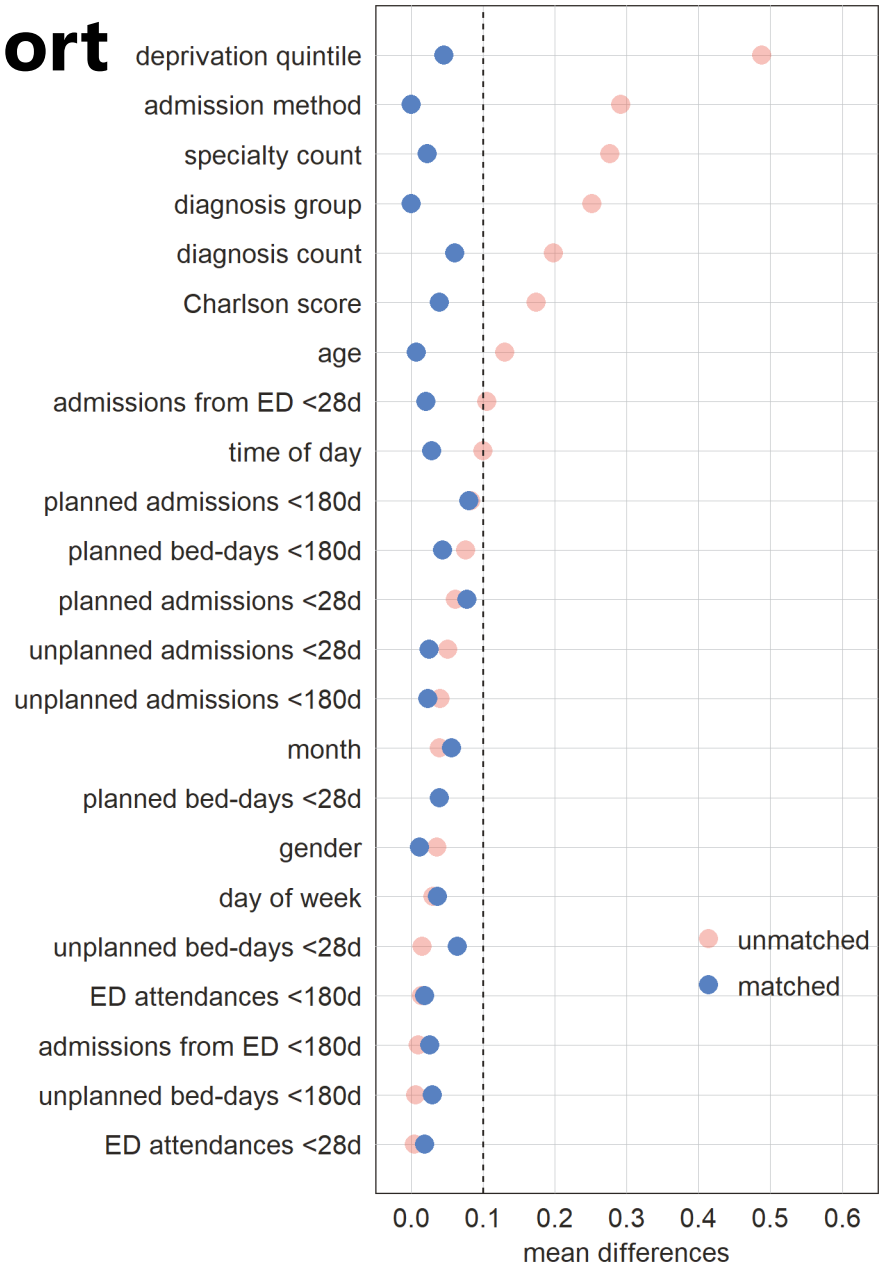
The potential for valid causal inference is dependent on achieving balance across the range of covariates used for matching. The plot right shows the assessment of balance—before and after matching—for RAID cases seen in ED. Balance was assessed by checking absolute values of the 'standardised mean difference'. A standardised difference of 0 indicates no difference between the groups; higher values indicate greater differences between the two groups. In general, standardized mean differences should be as close to zero as possible, but an upper limit of .1 is widely used as a conservative threshold for assessing imbalance that would lead to biased effect estimation (1). With the exception of month of attendance and day of week of attendance, after matching the SMD for all covariates was below the .1 threshold. Further information on the characteristics of the groups before and after matching is included in appendix A-1.



(1) Stuart EA et al. Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of Clinical Epidemiology*. 2013;66(8), S84–S90.e1.

Assessment of matching—inpatient cohort

The plot right shows the assessment of balance—before and after matching—for RAID cases seen in general wards. After matching the SMD for all covariates was below the .1 threshold. Further information on the characteristics of the groups before and after matching is included in appendix A-2.



Results—ED cohort

After matching, the matched comparison group and RAID patients group were exactly alike with respect to the distribution of patients by diagnostic group and arrival mode in ED. The two groups were similar across the full range of matching variables considered (see previous slides).

Outcome	Crude odds [1] RAID cases	Crude odds matched control patients	Relative difference from matched control group (adjusted odds ratio)	95% confidence interval [2]	p-value
Admission from ED	0.33	0.41	23.9% lower	12.3% lower to 33.9% lower	<0.001

The odds of admission from ED were lower for patients treated by the RAID team than for the matched comparison group—0.33 on average, compared with 0.41. After adjusting for any residual differences in the baseline characteristics of the two groups, the odds of admission for RAID patients were 23.9% lower than for patients in the matched comparison group. The 95% confidence interval suggested a relative difference in the range 12.3% to 33.9% lower.

[1] The odds of an event of interest occurring is equal to the probability of an event occurring divided by the probability of the event not occurring.

[2] A confidence interval shows some of the uncertainty in results. Although our best estimate is that RAID patients odds of admission were 24% lower than the matched comparison group, the true difference might lie in the interval from 12.3% to 33.9%. There is only a very small (5%) probability that the true difference is less than 12.3% or more than 33.9%.

Odds and relative risk

Odds and odd ratios are a widely reported measure of effect size when comparing a binary outcome between two groups, but are not easy to interpret and their use can lead to misunderstanding when communicating results. The odds of an event of interest occurring is equal to the probability of an event occurring divided by the probability of the event not occurring. For example, if 100 patients attend ED and 20 are admitted the odds of admission for the group is 20 over 80, or 0.25. For expressing probability, risk and relative risk, are more intuitive. Using the same example, risk of admission is calculated as 20 over 100, or 0.2. Odds ratios and relative risks (or risk ratios) are simply the ratio of the probability of an outcome (expressed as odds or risk) between two groups.

For patients treated by the RAID team, the 'adjusted odds ratio' for admission from ED was 0.76 (95% confidence interval 0.66 to 0.88). This means the 'odds' of admission in the RAID group were 23.9% lower than in the comparison group. It is important to understand that this finding is not the same as a 23.9% reduction in risk (or chance) of admission. When the outcome of interest is not rare in the population (and for most groups, chance of admission from ED is not rare), the odds ratio—if used as an estimate of relative risk—will overstate the effect of the treatment on the outcome measure.

Results—ED cohort (expressed as a relative risk)

In medical research, odds ratios are often the default choice for reporting the effect size comparing two groups in terms of an outcome that is either present or absent. This is because they are straightforward to estimate from statistical models which describe how the chance of an event occurring depends on a number of covariates or predictors. The relative risk (or risk ratio) is more intuitive, but cannot be easily obtained from the same statistical models.

To aid interpretation, we converted the odds ratio to a relative risk (1). On average, for patients treated by the RAID team their risk (or chance) of admission was 18.7% lower than for patients in the matched comparison group. The 95% confidence interval suggested a relative difference in the range 9.3% to 27.3% lower.

Outcome	Crude risk RAID cases	Crude risk matched control patients	Relative difference from matched control group (adjusted relative risk)	95% confidence interval	p-value
Admission from ED	0.25	0.29	18.7% lower	9.3% lower to 27.3% lower	n/a

(1) Grant RL. Converting an odds ratio to a range of plausible relative risks for better communication of research findings. *BMJ*. 2014;348: f7450.

Results—inpatient cohort

After matching, the matched comparison group and RAID patients group were exactly alike with respect to the distribution of patients by diagnostic group and method of admission. The two groups were similar across the full range of matching variables considered (see start of results section).

Outcome	Crude length-of-stay RAID cases, mean (sd)	Crude length-of-stay matched control patients, mean (sd)	Relative difference from matched control group (adjusted ratio)	95% confidence interval [1]	p-value
Average length of stay (days)	5.01	5.67	10.8% lower	9.2% lower to 12.4% lower	<0.001

The average length of stay was lower for patients treated by the RAID service than for the matched comparison group—5.01 days on average, compared with 5.67 days. After adjusting for any residual differences in the baseline characteristics of the two groups, length of stay for RAID patients was 10.8% lower than for patients in the matched comparison group. The 95% confidence interval suggested a relative difference in the range 9.2% to 12.4%.

[1] A confidence interval shows some of the uncertainty in results. Although our best estimate is that the average length of stay for RAID patients was 10.8% lower than the matched comparison group, the true difference might lie in the interval from 9.2% to 12.4%. There is only a very small (5%) probability that the true difference is less than 9.2% or more than 12.4%.

Potential for cost savings

Previous research has highlighted the potential for the RAID model to generate significant cost savings. The RAID model aims to reduce numbers of admissions from the emergency department and facilitate quicker discharge for patients from general wards. To provide an indication of potential cost savings, we have used our estimates of the effect of the RAID service to calculate numbers of admissions and bed days that might be avoided, and attached a cash value to this avoided activity. These indicative estimates of potential savings require careful consideration. The cost assumptions for admissions and bed-days are average unit costs from the NHS in England. These costs are fully absorbed costs so a reduction (or increase) at the margin in admissions or bed-days does not necessarily equate to the average cost. We have not considered the cost of providing the RAID service and this would need to be set alongside any savings.

Cost assumptions

Average cost of non-elective admission (1) [1] = £1500

Average cost of bed-day on general ward (1) [2] = £300

(1) NHS Improvement. Reference costs 2016/17: highlights, analysis and introduction to the data. Available from: https://improvement.nhs.uk/resources/-reference-costs/ [accessed 1st April 2018].

[1] The average unit cost of an excess bed day for the NHS in England in 2016-17 was £313. We were unable to source equivalent up-to-date costs for NI.

[2] The average unit cost of a non-elective inpatient episode (excluding excess bed days) for the NHS in England in 2016-17 was £1590. We were unable to source equivalent up-to-date costs for NI.

Potential savings—ED cohort

Savings calculations	
All RAID cases [1]	2650
Matched cases	2357
Admission rate—control group	0.29
Admissions—control group	682
Odds ratio (adjusted)	0.76
Admissions—RAID cases	558
Difference in admissions (cases – controls)	124
Est. savings (124 x £1500)	£186,625
Est. savings—inflated for missing data [2]	£209,824
Est. savings for 12 month period [3]	£132,520

[1] The following cases were excluded from the savings calculations: assessments where the patient was under 16 years, duplicate or inconsistent assessment records, multiple assessments associated with the same attendance, and assessments where there was no indication of a mental health diagnosis.

[2] Savings estimates were adjusted to account for missing data. This included assessment records missing a health and care number, records missing information required for matching, assessment records where a linked record in the ED dataset could not be located, and assessment records for which a suitable match could not be found.

[3] The study period was 19 months, the estimate was adjusted proportionally to a 12 month period.

Potential savings—inpatient cohort

Savings calculations	
All RAID cases (direct & influence groups) [1]	19,851
Matched cases	19,552
Mean length of stay—control group	5.67
Rate ratio (adjusted)	0.89
Mean length of stay—RAID cases	5.06
Total bed days—control group	110,870
Total bed days—RAID cases	98,889
Difference in bed days (cases – controls)	11,981
Est. savings (11,981 x £300)	£3,594,338
Est. savings—inflated for missing data [2]	£3,649,305
Est. savings for 12 month period [3]	£2,304,824

[1] The following cases were excluded from the savings calculations: assessments where the patient was under 16 years, duplicate or inconsistent assessment records, multiple assessments associated with the same attendance, assessments where there was no indication of a mental health diagnosis, and assessments with a learning disability diagnosis.

[2] Savings estimates were adjusted to account for missing data. This included assessment records missing a health and care number, records missing information required for matching, assessment records where a linked record in the ED dataset could not be located, records where length of stay was an outlier, and assessment records for which a suitable match could not be found.

[3] The study period was 19 months, the estimate was adjusted proportionally to a 12 month period.

Strengths & weaknesses

In a randomised experiment the assignment mechanism is known and ensures groups will only be randomly different from one another. Therefore, when using observational data it is important to understand the process by which patients are assigned to treatment. Hospital staff are encouraged to refer to the RAID service when they think a patient may have mental health and/or substance misuse needs. However, not all patients treated by the RAID service have a mental health diagnosis recorded in their care record, and not all patients with a mental health diagnosis are seen by RAID. Because we were concerned about our ability to replicate the assignment mechanism when selecting potential controls, and in particular the risk of systematic under-matching, we excluded all RAID cases with no mental health diagnosis indicated on their care record. For the RAID inpatient group we included all patients (admitted to either of the two intervention sites) with a mental health diagnosis, irrespective of whether they were treated by RAID. This was done because the RAID team provided training and support to hospital staff managing these patients during the study period.

One of the main threats to non-experimental studies is unobserved confounding. In our study, this could occur at hospital or patient level. We matched on a range of observed variables (such as age; arrival mode or method of admission; prior hospital utilisation etc.), but there may be differences between these groups that we did not observe (such as level of need, either at an individual or hospital level, or availability of mental health services outside the hospital—in the community) and that might have contributed to their outcomes.

Strengths & weaknesses

We selected comparison patients by examining diagnosis fields in hospital records for the presence of mental health related diagnoses. Although we used a single regional database for our analysis, the quality of the data and our results are dependent on the coding practices and accuracy of participating hospitals. Differences between hospitals in coding practice and accuracy could bias the results. Furthermore, there may be other differences in the characteristics of the participating hospitals that we were not able to allow for, such as differences in admission thresholds.

The use of a matched comparison group was intended to isolate the impact of the RAID service. We are not aware of any other major changes that occurred at the two intervention hospitals. However, outcomes may have been affected by other service changes occurring during the study period that we were not aware of.

Admission from ED and length of stay are primarily measures of resource use. Assessment of other potential impacts of the RAID service, such as health outcomes, patient satisfaction, staff confidence etc. was outside the scope of this study

Acknowledgements

This report was prepared by the Strategy Unit for the Northern Health and Social Care Trust. The Strategy Unit is a specialist team providing research, analysis and strategic change expertise from within the NHS.

The authors and funders would like to express their gratitude to the Belfast Health and Social Care Trust, South Eastern Health and Social Care Trust, Southern Health and Social Care Trust, and Western Health and Social Care Trust for their support and cooperation with this study.

We would also like to express our thanks to staff working for the Honest Broker Service (part of the Business Services Organisation) who assisted in the preparation of the datasets used for this study.

All statistical analyses were performed using R statistical software version 3.4.4 (1).

Errors or omissions remain the responsibility of the authors alone.

(1) R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017. https://www.R-project.org/.

RStudio Team. RStudio: Integrated Development Environment for R. Boston, MA: RStudio Inc.; 2016. http://www.rstudio.com/.

Appendices

A-1 Baseline patient characteristics before and after matching—ED cohort

	Potential controls	Matched controls	RAID cases
Total number of patients	23,463	2357	2357
Age, mean (sd)	42.54 (19.9)	36.65 (15.0)	37.53 (14.5)
Female	42.3%	43.7%	44.1%
Deprivation (IMD) quintile			
1 (most deprived)	53.2%	20.5%	20.7%
2	16.2%	27.4%	26.7%
3	11.6%	22.9%	21.1%
4	9.5%	18.7%	19.5%
5 (least deprived)	9.5%	10.5%	12.0%
Walk-in arrival [1]	51.4%	40.2%	40.2%
New attendance	99.1%	96.4%	95.8%
Incident type			
Self-harm	13.4%	18.2%	20.6%
RTA	0.4%	0.1%	0.2%
Other	86.2%	81.7%	79.2%

	Potential controls	Matched controls	RAID cases
Triage category			
Non-urgent	0.9%	0.3%	0.2%
Standard	13.0%	4.1%	4.2%
Urgent	53.5%	59.4%	58.5%
Very urgent	30.8%	35.1%	35.1%
Immediate	0.9%	0.2%	0.2%
Not known	0.9%	0.9%	1.8%
Diagnosis group [1]			
Self-harm	23.3%	45.5%	45.4%
Other substance abuse	9.1%	2.7%	2.7%
Alcohol	33.7%	10.1%	10.1%
Dementia	6.2%	0.1%	0.1%
Psychotic disorder	1.7%	3.5%	3.5%
Depression	20.4%	25.6%	25.6%
Undefined mental illness	5.6%	12.5%	12.5%

[1] Exact matching used.

Baseline patient characteristics before and after matching—ED cohort

	Potential controls	Matched controls	RAID cases
Prior utilisation, , mean (sd)			
Attendances prior 28 days	0.68 (1.71)	0.66 (1.5)	0.81 (1.8)
Attendances prior 180 days	2.98 (6.76)	2.77 (5.3)	3.10 (6.2)
Admissions from ED prior 28 days	0.16 (0.53)	0.18 (0.6)	0.21 (0.7)
Admissions from ED prior 180 days	0.74 (1.68)	0.76 (1.7)	0.83 (1.8)
Planned admissions prior 28 days	0.03 (0.34)	0.01 (0.3)	0.01 (0.3)
Planned admissions prior 180 days	0.15 (1.88)	0.08 (1.2)	0.09 (1.6)
Unplanned admissions prior 28 days	0.16 (0.51)	0.15 (0.5)	0.17 (0.5)
Unplanned admissions prior 180 days	0.76 (1.69)	0.65 (1.4)	0.66 (1.4)
Planned bed-days prior 28 days	0.02 (0.43)	0.00 (0.1)	0.00 (0.1)
Planned bed-days prior 180 days	0.24 (2.64)	0.05 (1.0)	0.14 (3.8)
Unplanned bed-days prior 28 days	0.27 (1.76)	0.18 (0.9)	0.22 (1.1)
Unplanned bed-days prior 180 days	1.97 (7.25)	0.95 (3.7)	1.19 (4.7)

	Potential controls	Matched controls	RAID cases
Month of year			
January	10.6%	10.9%	10.3%
February	10.0%	11.9%	10.9%
March	10.9%	14.3%	12.7%
April	5.2%	4.2%	5.7%
May	5.3%	5.6%	7.7%
June	5.3%	6.1%	6.2%
July	5.5%	4.0%	4.5%
August	5.7%	5.9%	5.4%
September	9.4%	6.4%	7.0%
October	11.0%	10.6%	9.5%
November	10.2%	9.4%	9.9%
December	10.8%	10.8%	10.2%

Baseline patient characteristics before and after matching—ED cohort

	Potential controls	Matched controls	RAID cases
Day of week			
Monday	14.0%	14.8%	14.1%
Tuesday	13.8%	15.6%	14.6%
Wednesday	13.7%	13.1%	15.1%
Thursday	13.1%	15.5%	14.2%
Friday	13.9%	10.8%	14.1%
Saturday	15.1%	14.1%	13.0%
Sunday	16.4%	16.0%	14.8%
Arrival time 07:00 to 19:00	45.0%	41.5%	40.8%

A-2 Baseline characteristics before and after matching—inpatient cohort

	Potential controls	Matched controls	RAID cases
Total number of patients	95,887	19,552	19,552
Age, mean (sd)	59.7 (19.6)	62.2 (19.7)	62.3 (19.7)
Female	52.6%	53.7%	54.4%
Deprivation (IMD) quintile			
1 (most deprived)	31.2%	14.7%	13.5%
2	22.4%	22.0%	22.9%
3	18.3%	27.6%	26.8%
4	15.0%	24.2%	25.0%
5 (least deprived)	13.1%	11.5%	11.9%
Method of admission [1]			
Elective	28.2%	22.7%	22.7%
ED	60.5%	71.6%	71.6%
Other emergency	5.7%	1.7%	1.7%
Maternity	1.1%	0.3%	0.3%
Transfer	4.5%	3.7%	3.7%

	Potential controls	Matched controls	RAID cases
Diagnosis group [1]			
Self-harm & poisoning	4.9%	6.3%	6.3%
Mental disorders – psychoactive subst.	37.5%	34.1%	34.1%
Dementia & Alzheimer's	15.1%	20.5%	20.5%
Other organic mental disorders	4.8%	4.4%	4.4%
Schizophrenia & delusional disorders	1.7%	1.3%	1.3%
Mood (affective) disorders	13.9%	17.0%	17.0%
Neurotic & behavioural disorders	6.3%	6.3%	6.3%
Other mental & behavioural disorders	2.2%	2.3%	2.3%
Medically unexplained symptoms	13.6%	7.7%	7.7%

[1] Exact matching used.

Baseline characteristics before and after matching—inpatient cohort

	Potential controls	Matched controls	RAID cases
Number of treatment specialties in spell			
1	84.1%	72.1%	73.1%
2	14.5%	24.6%	23.7%
3	1.3%	3.0%	2.9%
4+	0.1%	0.3%	0.3%
Number of diagnoses in spell			
1	4.8%	1.4%	1.9%
2	7.5%	8.5%	8.4%
3	9.9%	10.6%	10.5%
4	11.3%	13.0%	13.3%
5	11.2%	13.4%	12.9%
6	11.0%	13.7%	12.3%
7	10.1%	9.9%	10.4%
8	8.2%	7.5%	7.5%
9+	25.9%	22.0%	22.8%

	Potential controls	Matched controls	RAID cases
Charlson comorbidity score, index spell			
0	41.4%	35.4%	35.4%
1	23.4%	24.8%	23.8%
2	15.2%	15.6%	16.3%
3	9.0%	8.6%	8.8%
4	4.9%	8.7%	8.2%
5+	6.2%	6.9%	7.5%

Baseline characteristics before and after matching—inpatient cohort

	Potential controls	Matched controls	RAID cases
Prior utilisation, mean (sd)			
Attendances prior 28 days	0.31 (0.82)	0.33 (0.90)	0.31 (0.78)
Attendances prior 180 days	1.31 (2.93)	1.38 (2.93)	1.33 (2.52)
Admissions from ED prior 28 days	0.20 (0.59)	0.15 (0.44)	0.15 (0.45)
Admissions from ED prior 180 days	0.62 (1.34)	0.66 (1.33)	0.63 (1.33)
Planned admissions prior 28 days	0.89 (2.95)	0.83 (2.82)	1.06 (3.24)
Planned admissions prior 180 days	5.16 (17.82)	5.10 (17.52)	6.66 (20.78)
Planned bed-days prior 28 days	0.18 (0.48)	0.16 (0.43)	0.15 (0.43)
Planned bed-days prior 180 days	0.72 (1.43)	0.68 (1.30)	0.65 (1.31)
Unplanned admissions prior 28 days	0.30 (1.61)	0.29 (1.41)	0.35 (1.40)
Unplanned admissions prior 180 days	1.87 (6.40)	2.03 (7.22)	2.36 (7.23)
Unplanned bed-days prior 28 days	0.60 (2.40)	0.81 (3.19)	0.63 (2.37)
Unplanned bed-days prior 180 days	3.69 (9.51)	4.02 (9.93	3.75 (9.09)

	Potential controls	Matched controls	RAID cases
Month of year			
January	10.7%	10.4%	10.8%
February	9.9%	10.8%	10.7%
March	10.4%	11.3%	10.7%
April	5.4%	5.0%	5.2%
May	5.3%	4.9%	5.0%
June	5.4%	5.2%	5.2%
July	5.2%	6.4%	5.2%
August	5.7%	5.4%	5.3%
September	10.0%	9.7%	10.1%
October	11.0%	10.8%	11.0%
November	10.8%	9.9%	10.2%
December	10.3%	10.3%	10.5%

Baseline characteristics before and after matching—inpatient cohort

	Potential controls	Matched controls	RAID cases
Day of week			
Monday	16.5%	17.6%	16.4%
Tuesday	16.3%	15.8%	16.4%
Wednesday	16.1%	16.0%	15.7%
Thursday	16.1%	14.7%	15.4%
Friday	14.9%	15.4%	15.2%
Saturday	10.6%	10.7%	11.1%
Sunday	9.6%	9.8%	9.8%
Arrival time 07:00 to 19:00	61.3%	54.9%	56.4%