DETAILED PROJECT DESCRIPTION

Full title of project

Management by geographic area or management specialised by disorder? An evaluation of effects of an organisational intervention on secondary mental health care for common mental disorder.

Aims and objectives

The study aims to address the following research questions:

1. Is the quality of secondary mental health care for common mental disorder better when clinical teams are managed by geographically-based or product-based management divisions?
2. If so, how does this occur?

The study objectives are:

1. To perform a realistic evaluation of the introduction of a Mood, Anxiety and Personality Clinical Academic Group at South London and Maudsley NHS Foundation Trust, using methods of “logic modelling” (Soper et al. 2008) in order to design and interpret a set of time-series analyses using clinical data held in an anonymised repository of electronic patient records (the BRC Case Register). Time-series analyses will test (a) whether changes in single measures occurred in the expected direction at the expected time, (b) whether changes in one measure precede changes in another measure as expected and, where appropriate (c) whether changes are observed across diagnoses and between boroughs. Economic analyses will compare costs of care before and after the CAG program and relative to outcomes (cost-effectiveness).
2. To measure the inter-rater reliability of manual coding, where this is necessary in order to generate quantitative data from freetext data held in the BRC Case Register.
3. To disseminate findings locally and to potential research users more widely in the NHS as per the dissemination plan.

Background

Study Site: Population, Management Structure and Team Structure

Secondary mental health services for the London Boroughs of Croydon, Lambeth, Lewisham and Southwark are provided by South London and Maudsley NHS Foundation Trust (SLaM). In 2009, SLaM entered a partnership with Guy's and St. Thomas' NHS Foundation Trust, King's College Hospital NHS Foundation Trust and King's College London to form King's Health Partners, an Academic Health Sciences Centre (AHSC). The boroughs served by SLaM comprise a mixture of inner-city and suburban neighbourhoods. As is typical in the rest of London and also in other English cities, all electoral wards are in the upper deciles of estimated prevalence of common mental disorder (North East Public Health Observatory 2011; Heady & Ruddock 1996).

Before the study intervention (the CAG program - see below), SLaM's clinical teams for working age adults were managed through borough directorates. In each borough, the directorate was responsible for the care of adults with a range of psychotic and non-psychotic disorders. This structure of geographically-based divisions is typical of many NHS mental health trusts, which tend to have one adult mental health division per local authority area covered.

Part of the context of the CAG program is SLaM's team structure for general adult services. Like nearly all other mental health NHS Trusts, SLaM developed three types of functionalised team in response to the Mental Health Policy Implementation Guide (Department of Health 2001): Crisis Response / Home Treatment, Assertive Outreach and Early Intervention. These teams operated alongside typical generic CMHTs only in Croydon. In Lambeth, Southwark and Lewisham, most service users receiving secondary mental healthcare for common mental disorders have for several years been treated by eight short-term Assessment and Treatment (A&T) teams, with users with psychotic disorders tending to be managed by longer term Recovery and Support.
(R&S) teams as well as Early Intervention and Assertive Outreach teams; similar A&T and R&S teams will be introduced in Croydon in late 2011, replacing the existing generic CMHTs. Therefore, most clinical teams within SLaM’s borough services are already specialised in the care of one patient group.

This structure of separate A&T and R&S teams has a fairly long history and, allowing for terminological variation, it is used by many NHS mental health providers, although it remains less common than generic CMHTs. In the mid-1990s, we performed the PRiSM study, in which we established a mental health service with separate acute and continuing care teams in one area and compared its outcomes to an area in which a generic CMHT was established instead (Thornicroft, Straathdee, et al. 1998; Thornicroft, Wykes, et al. 1998). A survey in 1993 established that a division between acute and continuing care teams existed in around 10% of districts across England. National data are currently not collected, but a review of service directories on Trust websites indicates that 12 of 32 London Boroughs (38%) are served by similar bipartite teams. In the home counties, generic CMHTs are more common, but South Essex University Partnership NHS Foundation Trust is currently consulting on reconfiguring to separate teams for Primary Care Inreach, First Response and Specialist Care (http://www.sept.nhs.uk/Mental-Health/Outpatient-Redesign-Project.aspx), Berkshire Healthcare NHS Foundation Trust will shortly be introducing a Single Point of Access team (Dr Sohan Derasari, personal communication August 2011), and separate Access and Recovery teams already operate throughout the western half of the area served by Kent & Medway NHS & Social Care Partnership Trust (http://www.kmpt.nhs.uk/Adult-Mental-Health-Services.htm). In summary, many other NHS mental health providers have a similar specialised team structure to SLaM; to the extent that this structure is a precondition for the CAG program, it is already widely present; its presence even in providers in the home counties indicates that it may be more widely applicable.

There is little evidence regarding the relative effectiveness of the A&T / R&S structure versus generic CMHT, or indeed regarding the effectiveness of other kinds of functionalised team. The PRISM study referred to above (Thornicroft, Wykes, et al. 1998) found that patients in the sector served by the acute & continuing care teams had slightly greater increases in met needs, satisfaction and quality of life compared to patients in the sector served by the generic CMHT, but this was associated with costs that were 10% greater relative to the other sector (although overall costs declined in both sectors relative to the period before the introduction of any form of CMHT, mainly due to reduced bed use). Analysis of Mental Health Minimum Data Set data for England initially suggested that the introduction of Crisis Response / Home Treatment teams had resulted in reductions in the number of admissions, but no significant change in bed days, whereas the introduction of assertive outreach teams was associated with increased admissions and no significant difference in bed use (Glover et al. 2005; Glover et al. 2006). This lack of effect of assertive outreach corresponds to the mixed results of assertive community treatment and similar treatments in the UK (T. Burns, Creed, et al. 1999; T. Burns, Catty, et al. 2007). The effects of home treatment on admissions were significantly less than those reported by early implementers and early trials (Dean & Gadd 1990; J Hoult et al. 1984; Marks et al. 1994; Minghella et al. 1998; Harrison et al. 2001); notably, a reanalysis of the data on home treatment teams initially analysed by Glover et al has suggested that there was no effect on admissions (Jacobs & Barrenho 2011). Trial evidence of the efficacy of teams for early intervention in psychosis is inconclusive (M. Marshall & Rathbone 2011). We have previously suggested that different forms of team organisation are unlikely to have effects on outcomes unless they lead to greater use of effective treatments (Thornicroft et al. 1999). The existence of varied team structures where policy guidance is absent is of interest; research into the formation of different kinds teams in social work departments suggests that such arrangements (both patch working and specialised) arise largely organically out of pragmatic considerations (Challis & Ferlie 1986; Challis & Ferlie 1987; Challis & Ferlie 1988). While our study will not directly address the effectiveness of A&T/R&S teams versus generic CMHTs, we will ground it in a more detailed

1 These boroughs are Hammersmith & Fulham (West London Mental Health NHS Trust); Barnet, Enfield and Haringey (Barnet Enfield and Haringey Mental Health NHS Trust); Barking & Dagenham, Havering (North East London NHS Foundation Trust); Greenwich (Oxleas NHS Foundation Trust); the four boroughs served by SLaM; Richmond (South West London and St Georges NHS Trust; Mike Firn, SWLSG, personal communication Aug 2011)
appraisal of the current variety of team structures in mental health services, and our analyses
will take account of the varied history of team specialisation across SLaM.

Since 2009 there have been around 33,000 completed episodes of treatment within the A&T
teams serving Lambeth, Southwark and Lewisham and generic CMHTs in other parts of the Trust;
a diagnosis is recorded in structured data for 85% of these. Within Croydon’s generic CMHTs,
46% of episodes had a diagnosis of common mental disorder, while this was the case for 63% of
episodes in A&T teams, reflecting the greater concentration of common mental disorder in A&T
teams, even before the CAG program. Breaking down the case-mix of A&T teams since 2009 in
greater detail, depression is recorded in 38%, anxiety and related disorders in 18% and
personality disorder in 7%. A diagnosis of psychosis or bipolar disorder was recorded for 29%.

The CAG Program
At the time of its formation in 2009, King’s Health Partners adopted Clinical Academic Groups
(CAGs) led by a clinician, a manager and an academic as a uniform system of operational units.
Crucially, these relate to specific disorders or client groups. King’s Health Partners emphasises
the role of CAGs as a means to ensure that ‘the learning from research is used quickly, consistently
and systematically to improve clinical services’ (King’s Health Partners 2011) but did not dictate
how CAGs should be implemented as management structures within partner organisations.
Therefore, within the three NHS Trusts, CAGs have not assumed a consistent status. For King’s
College Hospital and Guy’s and St. Thomas’, CAGs exist in parallel with separate existing
management structures, although they draw on the services of many of the same clinical and
non-clinical managers. SLaM, however, used CAGs as the vehicle for a major management
reorganisation in which the existing management units were abolished and replaced by the CAGs
themselves, with CAGs assuming managerial responsibility during 2010. The greatest change was
the replacement of four borough directorates covering all adult mental health / general
psychiatric services with a Psychosis CAG and a Mood, Anxiety and Personality CAG (MAPCAG).
Therefore, the introduction of CAGs has led to a change in the practice of management broader
than just the inclusion of academics in the management of clinical services; rather, SLaM has
replaced one multidivisional structure of geographical divisions (G. Johnson et al. 2005, p.399)
with another composed of product-based divisions. Underneath, this management layer, clinical
teams have undergone only limited reorganisation, with most of the clinical teams previously
managed by borough directorates already being sufficiently specialised in the care of one patient
group that they were able to be simply adopted by one of the CAGs. The essential structural effect
of the CAG program is that clinical teams concerned with any branch of adult mental health /
general psychiatry are now led by a management unit responsible for the disorders that are the sole
or main focus of each team, rather than one that has a broad responsibility for adult mental health
services for a geographic sector. Our study focuses on common mental disorder (mainly
depression, anxiety disorders and personality disorder). Common mental disorder has the
Greatest impact of any illness: depression causes the most disability of any illness in the UK and
costs of depression and anxiety equate to 1% of GDP (Layard 2006). As outlined above, SLaM’s
services for common mental disorder are mainly provided through A&T teams; all of these are
now managed by the MAPCAG, along with several tertiary services for common mental disorder.

SLaM’s CAGs represent a kind of management specialisation complementary to team
specialisation. As set out below, we suggest that theoretical work on product line management
provides a preliminary theory of why such management specialisation might have effects on the
care and treatment of mental disorder, but overall, the CAG program lacks a clear program
theory, justifying our decision to perform the quantitative part of our evaluation within the
broader framework of realistic evaluation.

Service Line Management
We provisionally characterise CAGs in SLaM as a variant of service line management adapted to
an academic health sciences centre. Service line structures are certainly not exclusive to SLaM.
Among mental health NHS Trusts, they have also been adopted by Barnet, Enfield and Haringey
NHS Trust (http://www.beth-
mht.nhs.uk/about_us/trust_profile/our_position_in_the_nhs/servicelines? ts=1) and are being
implemented by Camden and Islington NHS Foundation Trust. More widely, they have been
extensively promoted in the NHS by Monitor, which has sponsored two national conferences on
Monitor has defined service lines as ‘the units from which the trust’s services are delivered, each with their own focus on particular medical conditions or procedures and their own specialist clinicians. Each unit also has clearly identified resources, including support services, staffing and finances’ (Monitor 2009, p.2). The conceptual basis for service or, as it was originally termed, product line management is to see a hospital as a “multiproduct firm”, with each product being a distinct, but relatively standard, bundle of goods and services produced for each patient through the use of a combination of intermediate outputs (e.g. an appendectomy produced by the application of nursing care, theatre time, etc.) (Fetter & J. L. Freeman 1986, p.42). Under this concept, the identification of product lines permits the use of clinical, activity and outcome data to work out the inputs applied in order to generate each product and therefore to analyse and control cost variation through a restructuring of how inputs are applied and combined. More broadly, service line management also ‘provides the means for administrators and physicians to understand and to be jointly responsible for the organization’s production process’ (Fetter & J. L. Freeman 1986, p.53).

Of course, however helpful such a schema may be in providing an initial orientation, exactly how it relates to the MAPCAG as implemented in SLaM remains to be determined. We aim to produce a model of the CAG’s effects derived from the explicit and tacit knowledge of those directly involved in the implementation, whether or not this resonates with the theory of product line management outlined above. Without beginning the process of logic model development it is difficult to know how such a model may look. However, “insider” clinician and academic applicants in this bid have suggested that they may see the CAG’s effects in terms of the application of concrete innovations intended to increase quality of care such as (a) the development of care pathways based on the Map of Medicine guidelines, (b) use of a structured measure to assess the presence of primary or co-morbid personality disorder, or (c) the use of methods to predict treatment resistance in depression2.

The promise of clinical engagement, better financial control, and improved quality of care is used to justify service line management (Monitor 2009), but empirical evidence of benefit of this managerial form of specialisation is thin, especially when contrasted to the evidence for specialised teams (which is not extensive—see above) and the evidence of improved outcomes for physicians, teams and providers specialised by dint of higher volume of care (Druss et al. 2004). We performed a scoping review in MEDLINE, finding 737 abstracts with a major subject heading of “Product Line Management”. Most of these studies are opinion pieces or contain very limited empirical data. A study of the introduction of three product lines at a Wisconsin hospital found increases in patient satisfaction, profit margin and efficiency (reduced LOS) over time (Turnipseed et al. 2007). However, there was not control group, and no attempt was made to ensure that these improvements were not simply attributable to an existing upward trend. A sophisticated evaluation of the introduction of a mental health and a primary care service line in Veterans’ Administration medical centers (Charns et al. 2001; Greenberg et al. 2003; Byrne et al. 2004) suggested a mixture of better, worse and unchanged performance. This however involved the introduction of a broad functional division that already exists, and in rather stronger form, in England, where it is now very unusual for a provider to have responsibility for both mental and general healthcare, and where the separation between primary and secondary is very clear. There is no reason to expect that the findings will also apply to narrower disorder-based service lines.

Realistic Evaluation and Logic Modelling
Most broadly, the CAG program may be considered as a restructuring from one form of multidivisional structure (geographically-based) to another (product-based) (G. Johnson et al. 2005, p.399); more specifically, we suggest, on the basis of theoretical work on product-line management, that this restructuring may permit a reorientation of management work to the

---

2 Not all of these innovations necessarily depend on managerial specialisation, and the individual effectiveness of at least one (care pathways) is uncertain (Evans-Lacko et al. 2008). However, it is conceivable that the managerial context is critical to their effects.
specific task of the care and treatment of particular disorders. But this does not really constitute a theory of the CAG program or any basis for planning or interpreting the results of an evaluation. The abstract nature of existing relevant theory contrasts strongly with the range of candidate understandings of the CAG program. Our response is to perform a "realistic" evaluation (Pawson & Tilley 1997). This will (a) be meaningful to the managers, clinicians, service users, families and GP commissioners who are involved in or affected by the MAPCAG implementation; (b) take advantage of their detailed contextual knowledge; (c) support replication by developing a textured and full account of exactly how the intervention had the effects that are demonstrated, and (d) increase research utilization by involving users in the evaluation (McKee et al. 2006).

According to Pawson & Tilley (1997), realistic evaluation "employs no one standard 'formula', other than the base strategy of producing a clear theory of program mechanisms, contexts and outcomes, and then using them to design the appropriate empirical measures and comparisons" (p. xv). Within our proposed evaluation, the task of "producing a clear theory of program mechanisms, contexts and outcomes" will mainly use the technique of logic modelling, whereas the role of the quantitative analyses which will occupy much of the project timetable will be to carry out the "appropriate empirical measures and comparisons". We propose a reciprocal relationship between these two work streams (see Methods below) with logic modelling directly influencing the planning of the time-series analysis, and the results of the time-series analyses feeding back into the later stages of logic modelling.

Pawson & Tilley (1997, chap.6) consider that the essential methodological imperative when using qualitative interviewing in realistic evaluation is to ensure that it is oriented specifically to the task of defining (and redefining) the context-mechanism-outcome relationships embodied in the program. Above all, "the researcher’s theory is the subject matter of the interview, and the subject (stakeholder) is there to confirm, falsify and, above all, to refine that theory" (Pawson & Tilley 1997, p.155). Logic modelling is a relatively new technique that allows this: we define it as the use of diagrammatic representations of key context-mechanism-outcome relationships within the setting of a qualitative interview in order to display the current understanding of the program and thereby prompt further data collection and refinement of the logic model. This use of "knowledge elicitation" techniques is similar to their use in applied forms of qualitative research such as rapid evaluation (Beebe 1995) or participatory rural appraisal (Chambers 1994). Logic modelling has recently been used in some important evaluation projects, most pertinently the evaluations of the Health Foundation’s Engaging with Quality Initiative and Engaging with Quality in Primary Care Initiative (Ling & Scoggins 2006; Ling et al. 2007; Ling et al. 2009; Ling et al. 2010; Soper et al. 2008).

In line with Pawson and Tilley’s directive, it encourage a formal focus on cause and effect and will enable us to develop a "narrative of change" that describes how those involved see their use of resources (money, authority, expertise, time, etc.) contributing to the intended benefits. It also helps to focus attention on the most important among what may be a large number of candidate context-mechanism-outcome relationships.

Participants in the logic modelling process will be those who would be expected to be able to contribute knowledge of the context-mechanism-outcomes of the program, or some part thereof: this includes managers and clinician-managers directly involved in the CAG program, clinicians within the CAG’s teams, as well as those GPs, service users and relatives who are the main “clients” of the CAGs services. Sampling will therefore be purposive, based on likely knowledge and understanding of the program and how it will work, from a variety of perspectives. Key “insider” respondents are already part of our project team (Tylee, Bindman, Davidson) and will advise on the selection of other respondents; recruitment of service users and relatives will be facilitated through the involvement of the PPI coordinator for the Mood, Anxiety and Personality CAG. As outlined in the proposal, we will use group interviews, which will be “focused” in so far as they will be oriented to the development of the logic model. We fully expect that different stakeholders will have different expertise in relation to parts of the program, and also that there may be disagreement concerning the relevant context-mechanism-outcomes: the aim of logic modelling is not to remove or combine such alternative or contradictory theories, but, if anything, to sharpen these in advance of a confrontation with the empirical data.

The basic logic model adapted from the Pawson and Tilley framework specifies a set of elements to include, namely inputs, activities, outputs, outcomes and impacts. In the logic models developed for the Engaging with Quality Initiative and Engaging with Quality in Primary Care
Initiative (Ling et al. 2010), these elements are linked in the above order, and each logic model enumerates mechanisms under each heading. The figure below shows the logic model for the QUEST project to improve mental health screening by school nurses (led by Prof Andre Tylee):

Logic models can be as complex as is needed. For example, the next figure depicts the logic model from an evaluation of an intervention to reduce illegal fly-tipping (courtesy of Dr Bryony Soper):

Although a qualitative methodology, logic modelling is well suited to evaluations in which
quantitative data are used to further explore, test and develop the model developed. These are often quantitative data that are already available. Therefore, in the Engaging with Quality Initiative, the teams’ own quantitative measurements made in the process of each quality improvement project were the foundation of the analyses performed by the evaluation team. In our own project we plan to support our evaluation mainly using data held in a database of anonymised electronic patient records—the BRC Case Register.

The BRC Case Register
South London and Maudsley NHS Foundation Trust (SLaM) has used a comprehensive electronic patient record—Patient Journey System (PJS)—for all clinical activity since 2006. Mental health has been one of the few care settings in England in which major progress has been made towards the implementation of electronic patient records (House of Commons Health Committee 2007). All mental health trusts in London now use similar systems. However, using these data for research and quality improvement poses challenges, in part because of the large quantity of patient identifiable and sensitive information that such records contain, but also because of the technical capacity required in order to manipulate and use data held in databases.

In 2007, SLaM and the Institute of Psychiatry, King’s College London successfully applied to NIHR to become the UK’s only specialist Biomedical Research Centre for Mental Health. A key part of the offer to NIHR was to set up a system for the use of anonymised electronic patient records for health research—the BRC Case Register, now operational since 2009 (Stewart et al. 2009). The Case Register contains all of the electronic patient records maintained by SLaM since 2006, but anonymised and optimised for data retrieval (querying). Studies now completed and either published or in press include studies of mortality (Chang et al. 2010) and life expectancy (Chang et al. 2011) among service users with serious mental illness, of mortality among service users with substance misuse diagnoses (Hayes et al. 2011), and of residential mobility (Tulloch, Fearon & David 2011a) and homelessness (Tulloch, Fearon & David 2011b) among acute psychiatric inpatients. BRC funding was recently renewed until 2017, and in recognition of the progress made in informatics, our BRC has assumed a national leading role for informatics in the BRC programme.

PJS contains a variety of structured data (numbers, dates or short standardised text) and unstructured data (free text), which together are equivalent to all of the data that would have been held across a computerised Patient Administration System and paper casenotes. Structured data include personal details, demographic details and details of clinical activity such as appointments, dates of periods of service by clinical teams and by the trust, and details associated with free text records—for example, the date on which a particular item of correspondence was created and added to the case notes. Free text data comprise mainly the aforementioned free text progress notes, and also correspondence. Anonymisation has two stages. In the first, patient identifiable information is stripped from the structured data in the database. Therefore, dates of birth are truncated to month and year of birth, ethnicity is grouped into broad categories, addresses are converted to the corresponding Office of National Statistics Output Area, and names of the service user and contacts are removed. A pseudonymous identifier (the BRC ID) is created, replacing trust and NHS IDs. In a second stage, free text data are cleaned of names: wherever names or recorded aliases for the user or his/her relatives are encountered in the free-text, they are replaced with ZZZZZ or similar. Once cleaned, data are organised into around 100 data tables ranging from the small (some tables containing infrequently used test scores) to the extremely large (the “Event” table containing free-text progress notes, which has over 10 million rows). The security procedures in place at the BRC have been reviewed by Oxfordshire REC and the BRC is treated for ethics purposes as an anonymised database: that is, access is granted after review of applications by a security committee, but formal ethics approval is not required.

Case Register data may be accessed through two interfaces. The most flexible and powerful interface is the direct interface to the SQL database, allowing data to be accessed and processed directly from standard statistical and data management software, such as Stata or SAS. SAS, unlike Stata, can accommodate free text data, and the BRC has invested substantially in a program of building capacity for the use of SAS aided by our partners, Amadeus Software Ltd. All BRC personnel have received training in SAS Enterprise Guide. Dr Alex Tulloch (co-applicant) has
attended training courses in SAS Programming, SAS Macro language and Advanced SAS Programming.

A major challenge in using Case Register data is how to generate numerical data suitable for quantitative analysis from the abundant free-text records, each of which may extend to 32,767 characters. The first step is to retrieve only records meeting specified criteria, for example those containing particular words or those which meet some criteria in other linked fields (for example, records made during a time interval defined separately for each of a sample of individuals). This can achieve important economy of effort. For example, in studies of homelessness and residential mobility among inpatients (Tulloch, Fearon & David 2011a; Tulloch, Fearon & David 2011b), 4485 admissions were selected according to defined criteria. It was then necessary to select only those records containing the terms “homeless”, “NFA” or “no fixed abode” which were also made in relation to each of the selected service users and during the admission recorded for each service user. (This required the use of SAS’s PROC SQL). This process alone selected around 2000 free text progress notes selected from the total of 10 million records, a process which takes a couple of minutes at most. However, to read 2000 such progress notes in their entirety remains a substantial task. This can be eased by using SAS to insert “tags” that change font colour (red) and weight (bold) for the target words when the data are displayed in Excel. The entire process of data extraction and coding then took no more than a day.

The present proposal envisages a similar but more intricate strategy in which a set of treatment episodes are defined using structured data (diagnoses and clinical activity), and then a defined selection of data is displayed for each selected episode, allowing simultaneous coding of a number of items (see “Methods”). Similar methods of “facilitated manual coding” have been used in a RAND project to develop a national system of performance indicators for the Veterans’ Administration’s mental health service (Watkins et al. 2011). To perform this coding, it will be necessary to display a spreadsheet of data for each treatment episode that must be coded, and then retain the coded data for analysis. Much of the required SAS programming has already been developed in a project to capture incident cases of psychosis in the Case Register (a BRC-funded collaboration with Amadeus Software Ltd): this programming can be readily and independently modified. This first-episode program has several steps. First, a SQL query retrieves a selection of data for individuals not already present on a cumulative database of first-episode psychotic patients and not already diagnosed as having a psychotic disorder, and whose recent free text entries contain particular words of interest such as “delusion” or “hallucination”. Second, these data are imported into SAS and then automatically outputted in a format suitable for manual coding. This involves splitting data into a multi-worksheet Excel workbook, such that each worksheet (tab) contains only data relating to a single person (in the case of our proposed project, each worksheet would similarly pertain to a single episode of care). Within each worksheet, data are arranged in a way that is conducive to rapid coding. A row of cells on the left of the worksheet comprises demographics and cells for the entry of coded data (whether the individual fits the study definition, the date at which this judgement was made, comments, etc.). A set of columns on the right of the worksheet contains the data for coding, which are drawn from multiple database tables. At the top, any previous diagnoses, and their dates, are displayed in date order; below that, any previous episodes of treatment are arranged, again in date order; and below that come the core data used in order to decide on the coded values: progress notes, correspondence and Health of the Nation Outcome Scales scores, again arranged in date order. As in the example above, targeted words are displayed in colour and in bold. The figure below is a screenshot of the current version of the Excel spreadsheet outputted by SAS:
Once each workbook has been manually processed, a third stage reads in the data and uses the new coded data to update the cumulative table referred to above. (In our proposal, instead of this upload stage, the tabulated data would be used directly for analysis).

Both projects described above selected free-text for manual review based on matching of one of several target words. However, additional efficiency is possible through the use of Perl regular expressions which permit more sophisticated pattern matching to be used to select and match free text (McGowan 2006). These techniques have been used with BRC data in order to select text recorded under headings such as “Presenting Complaint” and “History of Presenting Complaint”. We will use them to accommodate simple alternate phrasings or versions of phrases such as “seen for first session of CBT” and also to extract text including typical ICD-10 codes (F20.0, F32.2 etc.) from correspondence, where a diagnosis is typically entered in this form. By this means, data may be added to the store of structured data on diagnoses.

The project plans and costings presented in this research plan assume that only the techniques above will be used. However, work with academic partners at the University of Sheffield may yield more sophisticated strategies for extracting data from free text using text mining (Cunningham 2002). We have developed prototype utilities for extracting data on diagnosis, smoking status and antipsychotic prescribing data from free text; it may be that these will yield additional usable data over the proposed project timetable.

At the time of writing, preparatory work for the present application includes a program able to extract all activity from teams currently managed by the MAPCAG; teams which would but for name changes and reconfiguration have been managed by the MAPCAG, such as older A&T teams; and generic CMHTs. These data provide a complete record of the clinical activity of the trust. As noted above, there are around 33,000 such episodes since 2009. We have also calculated trends in data completeness for ICD-10 diagnosis, in order to assess the extent to which such data may be used for analysis, finding that a diagnosis was recorded for 85% over this time.
period. Because data completeness is lower before 2009, we have revised our project plan, and propose to analyse the effects of MAPCAG implementation from 2009 onwards, that is, with a one-year lead in to the year during which the MAPCAG assumed managerial responsibility.

Need
As outlined above, we provisionally characterise SLaM’s CAGs as a variant of service line management adapted to an academic health sciences centre. The organisation of secondary health care services on a service-line basis has an appealing rationale: By defining the products that a health care provider produces, and organising clinician-managers and managers around the delivery of those products, it may allow these managers to fully focus on the delivery of effective, efficient, timely, safe, equitable and acceptable care for specified disorders. Even in acute settings where some organisation of services on a disorder basis has long been present, the potential for improvement through the definition of more homogeneous health care products and increased managerial specialisation may not have been fully realised. In principle, the development of a PbR tariff may assist this, permitting the identification of cases with supposedly similar resource requirements and the revenue that they attract. In UK mental health services, which have been organised largely on a geographical basis, the introduction of management specialised by disorder would be a fundamental shift. Our study will show whether CAG implementation achieved the desired results in relation to one of the commonest disorders treated in secondary mental health care services, namely depression, or whether care was superior when delivered by management units also responsible for other mental health care. We will also show how any effects were achieved, assisting any attempt at successful replication.

“Pure” service line management has been extensively promoted by Monitor in the last few years through written reports (Monitor 2009), a series of toolkits and through two national conferences. It is already being implemented in NHS Trusts including mental health trusts such as Barnet, Enfield and Haringey NHS Trust and Camden and Islington NHS Foundation Trust. SLaM’s model of CAGs differs from typical service line management because an academic is involved in each management unit, but the essential features are otherwise identical: it would perhaps be most readily replicable in absolutely identical form in other academic health sciences centres and in NHS Trusts with associated CLAHRCs. Overall, this is not at all a “special” intervention, impossible to implement elsewhere.

Despite the level of interest in service line management, its evidence base is very limited. The King’s Fund is in the process of performing a (qualitative) case study of factors associated with outcomes of service line management implementation in seven NHS Trusts, and preliminary results were presented at the most recent Monitor conference in May 2011 (Foot 2011). This appears to be the only current study of service line management in the UK. A scoping review of Medline found 737 studies with “Product Line Management” as a major topic heading but most of these are opinion pieces or contain very limited empirical data. A study of the introduction of three product lines at a Wisconsin hospital found increases in patient satisfaction, profit margin and efficiency (reduced LOS) over time (Turnipseed et al. 2007). As outlined above, an evaluation of the impact of introduction of broader mental health and primary care divisions in VA medical centers is poorly applicable in the NHS (Charns et al. 2001; Greenberg et al. 2003; Byrne et al. 2004). There are also US studies of the existence (Eastaugh 2001; Gonzales 1997) or non-existence (Kass 1987; Kass 1998) of economies or diseconomies of scope, i.e. efficiency benefits of some forms of product specialisation at the provider level, and studies of the relationship between hospital competition and specialisation (Calem & Rizzo 1995): these studies, which undoubtedly address topics that are closely related to service line management, especially in the US, are again poorly applicable to NHS providers, which in general seek to find the optimal way of managing largely comprehensive services. In summary, there is little or no evidence regarding the consequences of the introduction of service line management in the NHS.

In order to be useful to the NHS, our study must maximise its generalisability. We have done so by (a) studying an NHS Trust that serves a typical English urban location (see “Background”) and (b) assessing effects across four previously separate geographic divisions with differing histories and configurations. We carefully considered the possibility of extending our study to a second site but this is unfortunately impractical: the data infrastructure required for this study has been developed over 5 years and simply does not exist elsewhere, in either the acute or mental health
sectors. We have extended our study to cover the whole range of common mental disorder (an extension that was requested by “insider” team members, for whom it fits more naturally with the MAPCAG’s activities, and which reflects a preliminary intimation that the logic model for MAPCAG implementation may not differentiate among different ICD-10 diagnoses). However, we rejected the addition of a second SLaM CAG as this would be twice as costly and probably over-ambitious.

Methods

Preliminary Phase (From award date to study commencement in April 2012)
During the preliminary phase of the study we will:
(a) appoint required study staff
(b) continue to identify study respondents for WP1 beyond those already identified (PPI coordinator for MAPCAG, Mr Steve Davidson, Dr Jonathan Bindman, Dr Andre Tylee)
(c) make research ethics and research governance applications in relation to WP1

Work Package 1: Logic Model Development (Months 1 to 4, 10 and 22)
WP1 will be co-led by an external expert in logic modelling for evaluation (Dr Bryony Soper) and by a service user researcher (Dr Diana Rose), and assisted by an additional researcher recruited from the pool of researchers employed by the Institute of Psychiatry’s Service User Research Enterprise. Details of logic modelling are provided in “Background” above. The respondent group will comprise at least 6 clinicians, managers and GPs involved in and/or impacted by the MAPCAG, 3 service users and 3 relatives of service users. Respondents in the clinician/manager/GP group will include the research team members indicated above. These “insider” team members have already identified a number of innovations of interest within the CAG—see “Background” above. We intend that service user respondents will include the patient-public involvement representative for the CAG, who will also be centrally involved in identifying other user and relative respondents.

In months 1 to 4, two user-only meetings, two meetings only for relatives of service users, and two professionals-only meeting will first be held. We believe that separating meetings in these early stages is important because, with the exception of the MAPCAG’s patient public involvement coordinator, the MAPCAG is likely to be largely unknown to users and relatives, whose interactions with the trust are almost exclusively through contact with clinical teams. It is also essential to ensure that potentially contrasting views of the context-mechanisms-outcomes originating from each stakeholder group are given sufficient space to develop at the initial stages, before being brought into a broader process involving users, relatives and health service personnel together.

These initial meetings will be followed by the first joint meeting, which will be scheduled towards the end of Months 1-4. During this meeting, preliminary logic models developed out of each separate stream will be presented and the data collected will be used in order to develop a preliminary logic model combining elements from each respondent subgroup.

There will be further joint meetings in Months 10 and 22, where the logic model will be refined in the light of WP2’s results. The lead for WP2 (Dr Tulloch) will provide input regarding data availability, especially during the early phase of WP1. Data collection will be performed by Dr Soper, Dr Rose and a further service user researcher. Data will be transcribed and analysed using a thematic analysis, assisted by the use of nVivo. Outputs at each stage of WP1 will include a diagram of the logic model for MAPCAG implementation (as per Figures 1 and 2 above), as well as prose reports. The diagram will be mapped to proposed quantitative data items and will be used for respondent validation and knowledge elicitation.

Work Package 2 – Time series analysis (Months 1 to 22)
Data for WP2 will come from the BRC Case Register, from staff and patient surveys and by content analysis of archive materials, especially minutes of borough directorate meetings and CAG board meetings.

Work for this work package will build on preparatory work detailed in “Background” above. We
will define all episodes of care for depression, anxiety and personality disorder from 2009 onwards using a combination of activity data (periods of treatment under teams covered by the MAPCAG, which would have been covered under the MAPCAG and generic CMHTs) and diagnoses. Diagnoses will be defined both by ICD-10 codes recorded as structured data, as well as ICD-10 codes retrieved from free text (see “Background” above regarding the use of SAS in order to do this, and a note on the possible use of full text-mining to extract further codes). We shall restrict the study to secondary rather than tertiary services—that is, to Assessment and Treatment Teams and generic CMHTs—and take a random sample of around 2000 cases (10 cases per weekly timepoint (see below).

In order to analyse our data, we will need to take each of these episodes and append values for the set of measures decided on during WP1: these will comprise outcomes of interest as well as mediators for those outcomes.

It should be carefully noted that we have not prejudged the indicators that will be selected for analysis of WP2 as the intention is that the analysis plan will be determined by the outcome of WP1. In assessing the feasibility of WP2, we have therefore considered a large range of measures that we feel may potentially be used, and have considered the practicalities of data set construction for each of these. In “Background” we have sought to emphasise the wide experience that we have of dealing with data from the Case Register and our flexibility in relation to data management and coding.

In practice, dataset construction will be done in two main ways. Where the required data can be derived from structured data (numbers, dates and standardised short strings of character data), we shall program all the required data manipulation steps in SAS and therefore automatically create a dataset with one row per treatment episode, and in which each row contains the derived values needed for the analysis. For measures which cannot be so derived, i.e. measures which require the manual coding of free text or whose coding requires the manual review of structured data together with other data, the number of steps required is greater. First, data need to be got into a form suitable for being manually reviewed and coded. Second, that coding needs to take place, and third, the coded data (only) need to be extracted and joined to the one-row-per-episode dataset described above.

Regarding the first of these strategies, much potentially important data (e.g. clinical activity including activity classified by professional, routine outcome measures such as the Health of the Nation Outcome Scales, deaths, diagnoses, demographic data) already exists in structured form. Therefore, for example, should waiting times be decided on as an outcome of interest, these could be automatically calculated from structured data recording appointment dates and dates of first referrals. Appointments with psychologists may be identified using a field that records the profession of the person seeing the service user.

Regarding the second of these strategies, several potentially important measurements, most notably guideline concordance and antidepressant prescriptions, will require manual coding. This can be greatly facilitated, as in the Veterans’ Administration (Watkins et al. 2011), by the generation of customised data displays in which the necessary free-text is selected, arranged and displayed in order to streamline coding. We will adapt existing SAS code in order to do this (see “Background”).

We will measure interrater reliability of this “facilitated” coding and its validity compared to full casenote review. In this sub-study, a sample of 100 treatment episodes will be coded both by the study RA and by other coders from SLaM’s Audit and Clinical Governance office. We shall also compare results based on coding from the “data display” and coding based on unselected notes. We shall calculate measures of agreement (kappa) and confidence intervals for these.

We may need to make use of data derived from content analysis of minutes of management meetings. For complex non-literal content, coding will be performed by the RA. For literal counts of word frequency, we shall process word-processed documents using SAS.

Cost-effectiveness will be measured by combining cost data and data on clinical outcomes.
selected for analysis; we will measure costs from the provider perspective by combining activity data (inpatient care, clinician contacts, medication) with appropriate unit costs (reference costs, BNF, PSSRU). Measures of effectiveness may include measures of change on routine outcome measures such as the Health of the Nation Outcome Scales (HoNOS; Wing et al. 1998). Economic analysis will compare cost-effectiveness (service cost divided by outcome change) pre- and post-CAG. Where HoNOS scores are missing we will use multiple imputation and use the imputed scores in sensitivity analyses. Potential variables included in the imputation will be demographic characteristics, previous HoNOS scores and past service history.

For the main time-series analyses, we will create a dataset of indicators from January 2009 to December 2013, aggregating data at 208 weekly timepoints. For each week, we will randomly select ten episodes in which treatment has been ongoing for six weeks or has been completed in that week. (That is, we shall make a random selection from largely or completed treated episodes). Therefore the overall sample size will be just over 2000. Ratings will be made on records made up to that timepoint.

Where the logic model predicts that the CAG program will lead to a change in a measure, we will perform a piece-wise regression of that indicator, allowing for autocorrelation, and test for the interaction between slope and time-period that would indicate an effect of the CAG program. A failure to allow for autocorrelation in time-series analysis can lead to underestimates of SE resulting in inflated t-statistics and hence higher false positive rates (Koutsoyiannis 2001).

Initially, we will check for the most common first order autoregressive (AR(1)) form of autocorrelation using Durbin-Watson’s test. Under AR(1) the autocorrelation coefficient ρ can be estimated using:

\[ \hat{\rho} = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2} \]

and a test of autocorrelation can be performed using the Durbin-Watson’s statistic given by:

\[ d = 2(1 - \hat{\rho}) \]

If found significant, the time series (y) and the corresponding predictors (x’s) need to be pre-whitened using the following transformation:

\[ \tilde{y}_{t+1} = y_{t+1} - \hat{\rho}y_t; \quad t > 1 \]
\[ \tilde{x}_{(t+1)j} = x_{(t+1)j} - \hat{\rho}x_{tj}; \quad t > 1, \quad j = 1, 2, \ldots, p \]

The transformed data will then satisfy the usual assumptions (free from autocorrelation) and the piece-wise regression model will be fitted to the transformed data. More complex autocorrelation structures such as ARMA(p,q) can also be addressed.

Where the logic model predicts that change in one measure will precede change in another, we will estimate the effects of lagged variables on other variables and perform tests of Granger causality (Gujarati 2002).

In the analysis of each time-series, a sample of 208 timepoints has >99% power (at alpha 0.05) to detect a time x slope interaction that has a medium effect size (0.15)\(^3\). We shall also test the hypothesis that the selected effects differ across the four boroughs, and, where appropriate, whether effects vary by diagnosis.

**Contribution to collective research effort and research utilisation**

---

\(^3\) Post hoc power calculation performed using G-Power’s module for ‘linear multiple regression: fixed model, \(r^2\) increase’, with one parameter being tested, and five entered into the model.
Institutional capacity
The Institute of Psychiatry at King’s College London (IoP) and South London and Maudsley NHS Foundation Trust (SlaM) have a long tradition of collaboration. This has been deepened through the formation of an academic health sciences centre (King’s Health Partners) by these two organisations and King’s College Hospital and Guy’s and St. Thomas’ NHS Foundation Trusts. The Health Service and Population Research Department at the IoP is committed to promote and conduct health services research at local, national and international level; to conduct studies that are both policy and practice relevant; and to foster an excellent research environment for researchers. We work collaboratively with service users, families and mental health charities as well as with international researchers and organisations, including public and private enterprises. Recently, King’s Health Partners has formed a new Health Policy and Evaluation Institute headed by Professor Graham Thornicroft. Members are drawn from across departments and schools of KCL and come from a range of disciplinary backgrounds. The present study fits well with the goals of the new institute which include increasing the volume and quality of research into health services at KCL and strengthening research utilization within the partner NHS Trusts; it will increase local institutional capacity to perform such research.

Technical research capacity
We will develop skills in the following domains:

(a) Realistic evaluation techniques. These are precisely intended to maximise the value of evaluations for managers and other research users by producing a “narrative of change” that recognises how the use of managerial and other resources contributes to improvement. Specifically, we will use “logic modelling” techniques which have been developed in order to facilitate such evaluations (Soper et al. 2008).

(b) The use of routinely available data for evaluation. We will address the challenge for NHS managers of turning data collected during the delivery of health services into information for the measurement of quality and processes of care. The IoP and South London and Maudsley are national leaders in the development of infrastructure and techniques for the use of data from electronic patient records: this study will be a cutting-edge use of such data for health services research.

(c) The use of quasi-experimental techniques of analysis, specifically the use of multiple time-series data. These techniques, imported from econometric analysis, are often the only practical way of evaluating organisational innovations. However, they are underused in health services research, reflecting their relative unfamiliarity compared with standard epidemiological study designs and the different models of causality used in analysis (Dowd 2010).

Dissemination
A systematic review of barriers and facilitators of use of evidence by health policy makers suggests that uptake of research is facilitated by personal contact, timely relevance of the research and the inclusion of summaries with policy recommendations (Innvaer et al. 2002). Our dissemination plans will be guided by these principles:

(1) we will identify the research users for whom our research is likely to be timely and relevant, such as those mental health NHS Trusts with similar team structure, those attempting in other ways to achieve knowledge transfer in health care organisations e.g. CLARHCS (Kislov et al. 2011) and other Academic Health Sciences Centres, those considering restructuring in response to the introduction of mental health Payment by Results (which is likely to reimburse Trusts on the basis of loose disorder-based pathways), those introducing service line management and those commissioning mental health services, especially those for common mental disorder which are provided not just by mental health NHS Trusts but also by the IAPT programme

(2) we will make initial contact with these potential research users, and then support personal contact in several ways, for example, recruitment to our external reference group, ad hoc presentations and meetings, and encouraging participation in a conference on “organising for quality in mental health services” which we will organise at the end of our study.

(3) our dissemination materials (the one year report to SDO, the full report to SDO, the research summary and our PowerPoint presentation slides) will include clear summary points and policy recommendations.
The details of our dissemination plans are as follows:
(a) Between funding and start of study, and with the assistance of Professor Helen Lester (Chair), we will recruit a ten-member external reference group; with the assistance of the external reference group, we will define a group of organisations for whom it appears that our research will have particular salience, especially organisations which have, or are moving towards, a similar clinical team structure.
(b) In months 0, 6, 12, 18 and 24, we will present to the external reference group. As well as receiving feedback on scientific aspects of the study, we shall ask the external reference group to consider whether there are additional steps that we should be taking to maximise the relevance of our research to external research users.
(c) In months 1 to 4, 10 and 22, we will be developing our logic model; however, this is not simply a process of data collection because collecting data on our program theory will require us to present to our respondents the evolving logic model and quantitative data available at that time. As our respondent group will include key clinicians, managers and other stakeholders, local dissemination will therefore be built into our study design.
(d) In month 12, we will report to SDO on the initial process of logic modelling and progress in developing and validating data management and coding techniques for the time series analysis.
(e) We will present our protocol and interim results (in 2013) and our final results (in 2014) to the SDO/HSRN conference and Mental Health Research Network conference—this will be approximately in months 12, 14, 24 and 26
(f) In months 22 – 24, we will prepare the final SDO report, summary and slides. In collaboration with Prof Ewan Ferlie and members of the external reference group these will be tailored to requirements of research community and managers respectively. We will prepare a separate briefing for service users and carers in collaboration with SURE and Rethink. We will assist SDO and SDO Network in making sure that are materials are accessible to potential research users.
(g) In month 24, we will organise a day conference on “Organising for Quality in Mental Health Services”. We will seek to enlist the support of other bodies such as Monitor, the Royal College of Psychiatrists Centre for Quality Improvement (CCQI) and the Health Service Confederation’s Mental Health Network. As well as presenting our results, we will develop a programme of short talks and seminars on related topics of mental health services management and improvement, and will include a keynote address by an internationally-recognised speaker. We will invite local stakeholders including clinicians, managers, users and carers, as well as target research users from around England and Wales. We will distribute dissemination materials.
(h) From month 6 onwards, we will prepare at least three articles in high-impact peer-reviewed journals: (i) an account of the logic model development; (ii) a validation study for the measures used in the quantitative study, and (iii) main results. We have assumed that one of these publications will be in a Biomed Central journal and have included publication fees in our costs.
(i) Throughout the project we will take any opportunities for ad hoc contact and meetings with potential research users.

Plan of investigation and timetable

See “Methods” above and attached Gantt chart.

Preliminary Phase (From award date to study commencement in April 2012)
(a) appoint required study staff
(b) continue to identify study respondents for WP1 beyond those already identified (PPI coordinator for MAPCAG, Mr Steve Davidson, Dr Jonathan Bindman, Dr Andre Tylee)
(c) make research ethics and research governance applications in relation to WP1

Work Package 1: Logic Model Development (Months 1 to 4, 10 and 22)

Work Package 2: Time Series Analysis (Months 1 to 22)

Final report writing (led by Dr Tulloch) (Months 22-24)

Approval by ethics committees
Use of the BRC Case Register is covered by a class ethical approval by Oxfordshire REC B. Our study will require review and approval by the BRC’s Oversight Committee, which includes Dr Rob Stewart (collaborator).

Once funding is arranged, we will use the six month lead-in to make an application for ethics and research governance approval for WP1. We do not anticipate any difficulties with this application, which concerns a low-risk study.

**Project management**

*Project Management*

The project will be led by the group of co-applicants under the direction of Professor Thornicroft and Professor Tylee. This group (Prof Thornicroft, Prof Tylee, Prof McCrone, Dr Soper, Dr Rose, Dr Tulloch) will meet monthly. Dr Tulloch will perform day-to-day project management and meet more regularly with other team members and the primary investigators. An external reference group will meet every six months led by Professor Helen Lester (University of Birmingham) to review progress and outputs of the study (we will have three face to face meetings and two video conferences). Dr Tulloch will be responsible for initial drafting of written outputs to SDO.

**Public users / public involvement**

Service users and relatives of service users will form part of the respondent group that will construct the "logic model" around which the evaluation will be planned. Initially, we will adopt a three-track approach with separate user, family and professional groups—the "logic model" suggested by service users will be combined with that produced by clinicians, managers, and commissioners only at points of genuine agreement. We will ensure that the logic model includes elements suggested by all sub-groups. Informal feedback from the Trustwide Involvement Group indicates that service users may be uncertain about how and why CAG implementation has occurred, so this project constitutes a valuable opportunity for patient and public involvement. Local dissemination events will involve users, families and professionals.

The MAPCAG’s patient public involvement co-ordinator will both be a respondent in WP1 (as she has special knowledge of the CAG) and will also help us to recruit and involve other users and family members.

Dr Diana Rose is a co-applicant and part of the team managing the research project. She will jointly lead WP1 and will have special responsibility within the project team for involving service users and their relatives and ensuring that the evaluation addresses their needs. Dr Rose is a Reader in User-led Research and co-director of the Service User Research Enterprise (SURE) within the Health Service and Population Research Department (IoP). SURE, which was formed in 2001, comprises service user researchers involved in all aspects of research (design, data collection, analysis, interpretation and dissemination). SURE staff are people who have used mental health services as well as having the relevant research skills. SURE’s aims are to increase service user involvement in all aspects of mental health research, and to change the focus of research towards issues important to service users. SURE has collaborated on over 20 large scale studies including cohort studies and RCTs, most of which have been with participants with psychosis. SURE has particular expertise in qualitative methods.

Through Dr John Larsen (Rethink), we will trial national dissemination materials with staff, users and relatives accessed through Rethink’s two IAPT projects.

Thus, patient and public involvement in this project has elements of all three of the INVOLVE levels of PPI in research - consultation, collaboration and user-led research.
References

Burns, T., Catty, J., et al., 2007. Use of intensive case management to reduce time in hospital in people with severe mental illness: systematic review and meta-regression. BMJ, 335(7615), p.336-.


WP2 does not require subject recruitment. Interdependencies between WP1 and WP2 are illustrated above. Flow diagram for WP1 is as follows: