BMJ Open

Identification of the delivery of cognitive behavioural therapy for psychosis (CBTp) Using Electronic Health Records and Open-Text Information in a Large Mental Health Case Register

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015297
Article Type:	Research
Date Submitted by the Author:	17-Jan-2017
Complete List of Authors:	Colling, craig; Kings College London (Institute of psychiatry, psychology and Neuroscience), Psychological medicine; South London and Maudsley NHS Fundation Trust Evans, Lauren; King's College London, Institute of Psychiatry; South London & Maudsley Foundation NHS Trust Broadbent, Matthew; South London and Maudsley NHS Fundation Trust; South London and Maudsley NHS Fundation Trust Chandran, David; King's College London, Institute of Psychiatry Craig, Thomas; King\'s College London (Institute of Psychiatry) kolliakou, Anna; King's College London, Institute of Psychiatry Stewart, Robert; King's College London, Institute of Psychiatry; South London and Maudsley NHS Fundation Trust Garety, Philippa; King's College London, Institute of Psychiatry
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Health informatics
Keywords:	cognitive behavioural therapy, CBT, CBT for psychosis, CBTp, electronic Health records, EHR

Identification of the delivery of cognitive behavioural therapy for psychosis (CBTp) Using Electronic Health Records and Open-Text Information in a Large Mental Health Case Register

Craig Colling[1, 2], Lauren Evans[1, 2], Matthew Broadbent[1, 2], David Chandran[1], Thomas J Craig[1], Anna Kolliakou[1], Robert Stewart[1, 2] and Philippa A Garety[1, 2]

[1] Kings College London (Institute of psychiatry, psychology and Neuroscience) London, SE5 8AF, UK.

[2] South London & Maudsley Foundation NHS Trust, London, SE5 8AZ, UK.

Text (exc. figures/tables): 4,016

Author contact information – Craig Colling, Biomedical research Centre, Mapother House, Maudsley Hospital, Denmark Hill, London SE5 8AQ , craig.colling@kcl.ac.uk and (020) 3228 8584.

Abstract

Objective: Our primary objective was to identify, with high precision and recall, CBT delivery for people with psychosis (CBTp) using a novel automated method in a large electronic health record database. We also examined what proportion of service users with a diagnosis of psychosis were recorded as having received CBTp within their episode of care; during defined time periods; compared with published audits; and whether demographic characteristics differentially predicted the receipt of CBTp.

Methods: Both free text using natural language processing (NLP) techniques and structured methods of identifying CBTp were combined to create a set of results to generate estimates of precision and recall. Using inclusion criteria from two published audits, we identified an anonymised cross sectional sample of 2,579 and 2,308 service users respectively with a case note diagnosis of schizophrenia or psychosis for further analysis.

Results: The method achieved precision of 95% and recall of 96%. Using the National Audit of Schizophrenia 2 criteria, 34.6% service users were identified as ever having received at least one session and 26.4% at least two sessions of CBTp; these are higher percentages than previously reported by manual audit of a sample for the same Trust. Receipt of CBTp was found to differ significantly by age, ethnicity, team type and diagnosis.

Conclusions: The methods presented here are an efficient method for examining delivery CBTp: this approach saves time, is likely to be more accurate than manual audits and provides more scope to monitor effectively the delivery of CBTp including equitable access across demographic/service factors.

Strengths

- A key strength of this study were the large sample and the innovative approaches adopted to identify CBTp delivery within the clinical record

- The ability to replicate the inclusion criteria of two previous audits also allowed us to contextualise the findings, and the large data set allowed access to data by year and also to examine clinical or demographic factors influencing delivery. Clearly there are also a large number of other variables in the EHR which can also be examined – relating to the service user characteristics, service delivery setting, therapist characteristics and aspects of therapy provision such as assessments; number of sessions; discontinuation and drop out and clinical outcomes.

Limitations

- A limitation of this study was that it took place in a single (albeit large) service provider; however, our results have identified themes that are consistent with other findings in relation to CBTp provision

- This approach does not provide an assessment of quality

Introduction

Background

Pharmacotherapy as monotherapy for people with a diagnosis of psychosis or schizophrenia is no longer regarded as optimal treatment. National guidelines in many countries, including NICE for England and Wales, recommend that psychological therapies, in particular cognitive behavioural therapy for psychosis (CBTp) and family intervention, should be offered; NICE makes the recommendation they are offered to all people with the diagnosis of schizophrenia and their carers.[1] However, repeatedly, within the UK service users, charities such as Rethink,[2] policy makers and audits [3] and [4] have reported that only a small proportion of people are accessing these treatments. For example the Schizophrenia Commission reported that only about 10% of service users access CBTp.[5] To address these concerns the Department of Health and NHS England are undertaking various initiatives, including the IAPT SMI (Improving Access to Psychological therapies for severe mental illness) [6] programme and the new Early Intervention Access and Waiting Time initiative [7] both of which aim to drive up access. However, one area of uncertainty which will limit evaluation of progress, is whether we do have accurate baseline estimates of current levels of provision. A recent national audit (NAS2) [8] taking a random sample of 100 service users with a diagnosis of schizophrenia or schizoaffective disorder in the community in each of 64 participating mental health trust or health boards in England and Wales concluded that there are significant gaps in the availability of CBTp and family interventions. For example, this manual case note audit found that Trusts reported that on average 39% of service users had been offered CBTp and 19% of service users had taken up CBTp. However there are grounds for thinking that the NAS2 audit might be inaccurate. The audit provided no definition or criteria for psychological therapy provision, asked whether a service user had *ever* been offered or received therapy, and was based on reports by consultant psychiatrists. The audit report noted that responses probably encompassed a broader set of interventions than covered by the NICE recommendations. In contrast a detailed manual survey of a random sample of 187 records, reported a very much lower rate of offers (6.9%) and delivery (5.3%) of CBTp,[9] employing expert reviews of reported therapy record content, within a one year period, in one large mental health Trust.

Manually conducted audits of case notes and electronic records, such as NAS2, requiring individual responses of health professionals, are a labour intensive way of establishing these data, limit the number of cases that can reasonably be investigated and are too cumbersome to use routinely as practical tools to monitor implementation in each trust. The national minimum data set [10] does not currently require interventions to be recorded, although this may change. Although in the South London and Maudsley NHS Foundation Trust (SLaM) a structured drop down record for psychological interventions in electronic records is available, there is concern that, as non-mandatory, it is incomplete.

In the current study we therefore sought to develop a method of using automated text based searches of clinical records using natural language processing (NLP) techniques, supplemented by structured fields, to investigate how much this might enhance our ability to provide accurate routine automatic data reports and analysis, and thus provide an efficient method of monitoring the implementation of psychological therapy provision, overcoming the limitations of manual case note audits. The decision to focus initially on CBTp delivery instead of CBTp offer was a pragmatic one based on the perceived complexity and the resultant time required for each project.

This paper is focussed on the provision of CBTp within the UK but the challenges associated with implementation of CBTp for service users with psychosis are of international concern and relevance.[11]

Methods

Setting

SLaM is a large provider of mental healthcare, serving a catchment of around 1.3 million residents in four boroughs of South London (Croydon, Lambeth, Lewisham and Southwark). The majority of people with a diagnosis of a schizophrenia spectrum disorder are served by early intervention teams for the first three years from initial presentation and by promoting recovery teams subsequently.

Study design

Source of clinical data

The data for this study were obtained from the SLaM Biomedical Research Centre (BRC) Case Register and its Clinical Record Interactive Search (CRIS) application,[12] which accesses anonymised data from the electronic health records (EHR) of individuals who have previously received or are currently receiving mental healthcare from SLaM within a robust, service user-led governance framework.[13] At the time of writing this is over 265,000 service user records. We used CRIS to replicate the inclusion criteria for NAS2 and Haddock et al (2014) as means of comparison with two published audits. The SLaM BRC Case Register contains structured fields, such as those coding demographic information, as well as unstructured (but de-identified) free text fields from case notes and correspondence where history, mental state examination, diagnostic formulation and management plan are primarily recorded. The CRIS data resource has been approved for secondary analysis by the Oxfordshire Research Ethics Committee,[14] as well as a service user-led oversight committee which considers all proposed research before access to the anonymised data is permitted. The electronic health record system was implemented in SLaM services in April 2006 so provides an extensive repository of data for this study.

Overview of methodology

The initial step was to identify the delivery of CBT across all patient groups not distinguishing by diagnostic groups or other characteristics and then subsequently, and as the specific focus of this study,

to test the performance of the application for a sample of service users with a diagnosis of psychosis (CBTp).

Identification of Cognitive behavioural Therapy (CBT) delivery using CRIS

Natural language processing (NLP) techniques [15] were used to identify CBT delivery from free text fields within the BRC Case Register. The annotation strategy to identify whether a clinical record was a session of CBT was developed by three human annotators (CC, LE and MB) who also completed the initial feasibility which was signed off by an expert clinical lead (PAG). All annotations were double annotated by two human annotators, and disagreements were resolved by consensus and liaising with the clinical lead if required. Inter-annotator agreement was evaluated following each batch of annotations completed and the annotation strategy was updated according to issues raised and clarifications identified. Two annotators reviewed a training set of 300 instances in the development phase before annotating a gold standard dataset of 200 where the term "CBT" (or variants of) occurred and annotated as to whether the sentence that contained the term "CBT" was an actual session of CBT rather than a historic reference to therapy, referrals for CBT, decision not to offer CBT, or other reference to CBT which was not a therapy session. When a positive instance of CBT delivery was identified, the following features were recorded: session number, stage of treatment, the recipient of treatment and whether the CBT was delivered individually or via a group. Once the human annotations were complete, the training set was reviewed by the NLP developer (DC) to establish the rules to determine whether the CBT text is an actual session or not. These rules were coded using General Architecture Text Engineering (GATE) software.[16] Within the development process the impact of the rules applied to the training set were measured by the precision (equivalent to positive predictive value) and recall (equivalent to sensitivity). There is an inherent trade-off between the precision, and recall as one increase the other reduces, so there is a balance between what is more important in relation to the problem domain. We concluded that for this study an evenly weighted solution was preferred with a slight preference to precision. When precision is prioritised, this results in false positives being minimised which increases the confidence in the test to

correctly identify the positive outcome at the expense of incorrectly classifying some positive instances as false negatives. To illustrate this, Fig 1 shows the relationship between precision and recall in the classification task.

When all the rules were developed based on the training set, the model was tested against an independent gold standard dataset to evaluate how well the model performed on unseen data using precision and recall as the metric of evaluation. Once we were happy with the precision and recall on the gold standard, the resulting application was applied against the CRIS database and we further tested whether other relevant variables such as the professional group of the clinician who entered the clinical note, whether the clinical note was classified as a psychological therapy in structured data drop down menu, or whether the positioning of the CBT reference in the clinical document could be used to improve the performance of the application.

Identification of Cognitive behavioural Therapy for Psychosis (CBTp) delivery using CRIS

The output of the CBT application was tested against a sample of service users with a current diagnosis of psychosis to evaluate whether the precision and recall were of an acceptable standard or whether a specific CBTp application would need to be developed psychosis.

Within SLaM, psychological interventions can be recorded through a drop down box within the clinical record, but as a non-mandatory field the recording is potentially poor. To assess the quality and use of this field a senior clinician completed an audit of 100 documents where CBT was indicated within the drop down box, identifying whether the text associated with the document could be confirmed as a session of CBT.

Both free text and structured methods of identifying CBT were combined to create a single set of results which was used for analysis purposes. As the focus of this paper is to identify the delivery of CBT for patients with a diagnosis of psychosis the term CBTp is used from this point forward.

Research Question

The primary research question of the study was whether we could identify, with sufficiently high precision and recall, CBTp delivery using free text and structured methods in a large electronic service user record database, thereby improving feasibility, scope, scale and confidence of such an audit. We also examined how many and what proportion of service users according to inclusion and exclusion criteria employed in published audits (see below), with a case note diagnosis of schizophrenia or psychosis were recorded as having received CBTp within their episode of care using the CRIS database, during defined time periods, combining NLP and structured records. We then compared these data with the results of two published audits. Finally we examined whether demographic characteristics differentially predicted the receipt of CBTp.

Participants

We used the CRIS database to generate two large participant samples in this study, one replicating the inclusion criteria and the sampling time frame employed by the NAS2 audit and a second which replicated the Haddock et al (2014) audit inclusion criteria, allowing a comparison with each study.

1. **NAS2 audit inclusion criteria**

All individuals active for at least 12 months on 01/07/2013 aged over 18 in either early intervention or promoting recovery service with a recorded diagnosis of schizophrenia (F20.0 – F20.9) or schizoaffective disorder (F25.0 – F25.9). The NAS2 audit requested whether CBTp was "taken up" and we examined this in two ways: service users with at least one session of CBTp and service users with at least two sessions of CBTp prior to the census date.

2. **Haddock et al (2014) audit inclusion criteria**

All individuals active between 01/07/2012 and 01/07/2013 aged over 18 in either early intervention or promoting recovery service with a recorded diagnosis of schizophrenia spectrum diagnosis

(schizophrenia, schizoaffective, schizotypal and delusional disorders (F20.0 – F29.9)). CBTp delivery was defined as at least one session of CBTp within the 12 month audit period.

Demographic and service variables

The following variables were extracted for analyses: age, diagnosis, ethnicity, gender, and marital status. All data obtained were the most recent prior to the census date. Ethnicity was recorded according to categories defined by the UK Office for National Statistics and categorised for analysis purposes into three groups: black, other and white. Diagnosis is routinely recorded in clinical services using the ICD-10 classification system in drop-down fields, and was limited to schizophrenia spectrum (F20 – F29), with an additional sub-grouping applied in line with the NAS2 diagnostic categories of schizophrenia (F20.0 – F20.9), schizoaffective disorder (F25.0 – F25.9) and 'other schizophrenia spectrum' (F21, F22.0 – F22.9, F23.0 – F23.9, F24, F28 and F29). We used the NAS2 inclusion criteria to investigate the delivery of CBTp across: age group, diagnosis, gender, ethnic group and the service type.

Statistical Analysis

Descriptive statistics for demographic variables are reported as means and standard deviations for continuous variables (age at referral) and as frequencies and percentages for all other variables. A binary logistic regression model was used to examine the differences for proportions of cases who received CBTp and those who did not. We initially undertook an unadjusted analysis by age group, diagnosis, ethnicity, gender, marital status and service type to establish whether the receipt of CBTp differed by these demographic factors. We subsequently undertook a multivariable analysis, adjusting for potential confounders by including covariates (age, diagnosis, ethnicity, gender, marital status and service type) in the model except the variable of interest to explore whether the significant relationship identified within the unadjusted analysis remained after adjustment for potential confounders within the adjusted model.

Results

Precision and recall of identification of CBT in CRIS

Within the developed NLP CBT delivery application, the performance was validated against the independent gold standard resulting in precision (positive predictive value) and recall (sensitivity) for CBT annotations of 85% and 86% respectively. Following the development of the CBT NLP application, we concluded the precision would be improved with a tolerable reduction in recall if we applied the following post processing rule; to exclude CBT sentences that commenced after the first 200 characters of the clinical document. This post processing rule resulted in improving the overall performance of the application with an increase in precision of 12% to 97% and a reduction in recall 4% to 82%. The evaluation of the structured CBT entry alone resulted in a precision level of 89%. We then combined both methods, a measure was adopted to establish the recall of the combined method by reviewing the false negatives (FN) from the NLP app and examining whether they were identified by the structured method: of the 12 FN's identified by the NLP app 75% (9/12) were correctly identified by the structured data with the effect of increasing the recall from 82% (56/68) within the NLP app to 96% (65/68) within the combined method. By combining methods, we therefore achieved precision of 97% and recall of 96%. The NLP app resulted in identifying 26% additional service users who received CBT not recorded by the drop down box.

Precision and recall of identification of CBTp in CRIS

We further evaluated the developed NLP CBT delivery application against a sample of service users with a diagnosis of psychosis. The performance against the independent gold standard resulted in precision and recall for CBTp annotations of 81% and 85% respectively. We concluded the precision and recall would be improved if we applied the following post processing rule; to exclude CBTp sentences that commenced after the first 200 characters of the clinical document. This post processing rule resulted in improving the overall performance of the application for CBTp with an increase in precision of 14% to 95% and a reduction in recall of 7% to 78%. The evaluation of the structured CBT entry alone resulted in a precision

level of 89%. We then combined both methods, a measure was adopted to establish the recall of the combined method by reviewing the false negatives (FN) from the NLP app and examining whether they were identified by the structured method: of the 10 FN's identified by the NLP app 80% (8/10) were correctly identified by the structured data with the effect of increasing the recall from 78% (36/46) within the NLP app to 96% (44/46) within the combined method. By combining methods, we therefore achieved precision of 95% and recall of 96%. The NLP app resulted in identifying 21% additional service users who received CBTp not recorded by the drop down box.

Delivery of CBTp using sample based on NAS2 inclusion criteria

2,308 service users were identified in the dataset as fulfilling the NAS2 inclusion criteria. Service users had a mean age of 40.7 at referral (SD - 12.1 and range 18 – 83), and 1,806 participants had a diagnosis of schizophrenia ((F20) and 502 a diagnosis of schizoaffective disorder. The SLaM return for the actual NAS2 Audit was that 20% of the random sample of N=100 were identified as ***having ever received*** CBTp. In contrast, using the current method, 34.6% (799/2,308) were identified as having at least one session and 26.4% (610/2,308) were identified as having at least two sessions of CBTp. A breakdown of CBTp delivery by diagnostic group can be viewed in **Table 1 –** CBTp delivery by diagnostic groups using NAS2 audit criteria below.

Diagnostic group	n	% episodes with at least 1 CBTp session	% episodes with at least 2 CBTp sessions
Schizoaffective disorder (F25.0 - F25.9)	502	42.4%	32.9%
Schizophrenia (F20.0 - F20.9)	1,806	32.4%	24.6%
Total	**2,308**	**34.6%**	**26.4%**

We also explored the level of CBTp provision by year which can be viewed in Fig 2.

Delivery of CBTp using sample based on Haddock et al inclusion criteria

2,579 service users fulfilled the inclusion criteria within the same 12 month audit period. Service users had a mean age of 40.0 at referral (SD - 12.4; range 18 – 83) within the sample. We found that 12.8% (330/2,579) received CBTp interventions within the same twelve month audit period, whereas Haddock et al reported 5.3% in their sample.

We also examined a more recent time period: 2,597 service users fulfilled the inclusion criteria within a 12 month audit period within 2015. Service users had a mean age of 39.6 at referral (SD – 12.7; range 18 – 85) within the sample. We found that 14.8% (385/2,597) received CBTp interventions within the twelve month audit period.

Demographic predictors of at least one session of CBTp

The demographic characteristics of service users who received CBTp were compared with those who did not using our largest sample of N=2,579, which employed the Haddock inclusion criteria. The receipt of CBTp was more common in the following groups of service users – the younger group when compared with the older group, the White group when compared with the Black group, the schizoaffective disorder group when compared with the schizophrenia group and the Early Intervention for psychosis teams (EI) when compared with the promoting recovery teams (PR). Table 2 provides a summary of the unadjusted and adjusted multivariable logistic regression for receipt of CBTp by age group, diagnostic group, ethnic group, gender, marital status and service type.

CBTp delivery within a mental healthcare setting

Table 2 – unadjusted and adjusted logistic regressions for predictors of at least one session of CBTp

			Unadjusted			Partially adjusted*			Fully adjusted**		
Group	**n**	**%**	**Odds ratio**	**Confidence interval**	**Significance**	**Odds ratio**	**Confidence interval**	**Significance**	**Odds ratio**	**Confidence interval**	**Significance**
Age											
Under 41	1,346		1.57	1.24 – 1.99	<.001	1.57	1.23 – 2.01	<.001	1.32	1.01 – 1.72	.043
41 and over	1,233		Reference category								
Ethnic group											
Black	1,314		Reference category								
White	908		1.34	1.04 – 1.72	.024	1.40	1.08 – 1.80	.011	1.43	1.10 – 1.85	.007
Other	357		1.35	0.96 – 1.90	.081	1.33	0.94 – 1.88	.106	1.31	0.93 – 1.86	.122
Diagnostic group											
Other schizophrenia spectrum	271		2.26	1.63 – 3.14	<.001	2.02	1.45 – 2.82	<.001	1.52	1.05 – 2.20	.025
Schizoaffective disorder	502		1.53	1.15 – 2.03	.003	1.47	1.10 – 1.97	.009	1.48	1.11 – 1.98	.008
Schizophrenia	1,806		Reference category								
Gender											
Male	1,555		Reference category								
Female	1,024		1.15	0.91 – 1.46	.230	1.19	0.94 - 1.52	.155	1.20	0.94 – 1.54	.139
Marital status											
Single/divorced	2,339		Reference category								
Married/co-habiting	240		0.93	0.62 – 1.40	.729	0.90	0.60 – 1.37	.623	0.95	0.63 – 1.44	.809
Service											
Early Intervention	327		2.49	1.87 – 3.31	<.001	N/A			1.98	1.40 – 2.81	<.001
Promoting recovery	2,252		Reference category								
***Within the partially adjusted model the results were adjusted for age, ethnic group, diagnostic group, gender and marital status**											
****Within the fully adjusted model the results were adjusted for age, ethnic group, diagnostic group, gender, marital status and service**											

Discussion

To our knowledge this is the first attempt at using NLP techniques on free text entries, supplementing structured fields, in order to identify the delivery of one type of psychological therapy in a large health record data set. This was successful: we achieved a high level of precision (95%) and of recall (96%) which is consistent with other published CRIS NLP applications, which have measured other clinical activities or characteristics such as prescribed medication,[17] Mini-Mental State Examination score,[18] diagnosis [19] and service user characteristics such as, smoking status [20] and whether the service user lived alone. [21] The methods presented here are therefore an effective and efficient method to examine the delivery of CBTp: this approach provides significant time saving benefits compared with manual audits which require clinicians to complete a case note audit on an annual basis, while also providing a much more comprehensive and accurate overview of the delivery of CBT across all cases.

When using this method, we identified higher levels of CBTp delivery than previously reported in the SLaM contribution to the NAS2 audit using the same sampling criteria but a very much larger sample. We also found higher levels of CBTp delivery (about double) than that reported by Haddock et al (2014) in the same time period, albeit in a different service setting. This suggests that manual audits may result in under-reporting, presumably because of the limitations of clinician knowledge or readily accessible recording in health records, and our development is encouraging because it may result in both better quality output and much less time-intensive data collection. It is notable that the NAS2 audit enquired whether CBTp had ever been provided: the methods described here can search by year, which is clinically more useful; the data also might suggest that clinicians in responding to such an audit are typically considering perhaps the previous two years. Furthermore, when we conducted the sampling twice for 2013 and 2015 we found some evidence of a modest increase in provision - from 12.8% to 14.8%. However our results also continue to show that CBTp delivery falls very far short of the NICE recommendations of universal access. It is a matter of additional importance and concern that there do appear to be demographic predictors, suggesting access is inequitable in terms of age, diagnosis and

ethnicity. Improving access to psychological therapies can be enhanced by examining data such as these and targeting provision towards under-served groups.

Strengths

A key strength of this study were the large sample and the innovative approaches adopted to identify CBTp delivery within the clinical record. The ability to replicate the inclusion criteria of two previous audits also allowed us to contextualise the findings, and the large data set allowed access to data by year and also to examine clinical or demographic factors influencing delivery. Clearly there are also a large number of other variables in the EHR which can also be examined – relating to the service user characteristics, service delivery setting, therapist characteristics and aspects of therapy provision such as assessments; number of sessions; discontinuation and drop out and clinical outcomes.

Limitations

A limitation of this study was that it took place in a single (albeit large) service provider; however, our results have identified themes that are consistent with other findings in relation to CBTp provision and could indicate generalisabity but would warrant further investigation. This approach does not provide an assessment of quality.

Next steps

The opportunity provided by employing methods shown here allows the proactive analysis of large EHR-derived data sets. In the future, a refinement could be to identify CBT delivery data by using data from NLP and structured fields to identify a course of CBT treatment. Initial definitions regarding the development of a course of treatment would require at least two CBT sessions with less than a three month break between sessions and in addition utilising other NLP features such as the CBT session number and stage of therapy to enhance the creation of such a construct. We are also now working on

developing an application that identifies the delivery of other therapy types and applications which more precisely characterise the pathway from CBT being considered, through its offer and to actual receipt.

Contributors

Contributors PAG and RS conceived the study and manuscript. CC, LE and MB conceptualized the development plan. DC undertook the development. CC and LE tested the application during the development process. MB supervised the development process. CC undertook the data management and conducted the main analyses. CC and PAG led the drafting of the manuscript. AK contributed to the interpretation and presentation of the data. TJC, PAG and RS provided critical revisions to the content of the manuscript.

Competing interests

No competing interests.

Funding

This work was supported by the Clinical Records Interactive Search (CRIS) system funded and developed by the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London and a joint infrastructure grant from Guy's and St Thomas' Charity and the Maudsley Charity (grant number BRC-2011-10035). PG and TJC acknowledge BRC support. CC and LE are funded by SLAM. All other authors receive salary support from the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The above funding bodies had no role in the study design; in the collection, analysis and interpretation of the data; in the writing of the report; and in the decision to submit the paper for publication.

Data sharing statement

No additional data are available.

References

1 National institute clinical excellence. (2014). *Psychosis and schizophrenia in adults: prevention and management.* UK, National Collaborating Centre for Mental Health.

2 Rethink. (2008). *Your treatment, your choice survey: final report*. UK, Rethink.

3 Prytys M, Garety PA, Jolley S, Onwumere J, Craig T. (2011). Implementing the NICE guideline for schizophrenia recommendations for psychological therapies: A qualitative analysis of the attitudes of CMHT staff. *Clin Psychol Psychothera*, 18, 48–59. DOI: 10.1002/cpp.691.

4 Krupnik Y, Pilling S, Killaspy H, Dalton J. (2005). A study of family contact with clients and staff of community mental health teams. *The Psychiatrist*, 29(5), 174-176. DOI: 10.1192/pb.29.5.174

5 Schizophrenia commission. (2012). *An abandoned illness*. UK, Schizophrenia commission.

6 Department of Health. (2011). *Talking therapies: A four-year plan of action*. UK, COI.

7 Department of Health. (2014). *Achieving Better Access to Mental Health Services by 2020*. UK, COI.

8 Royal College of Psychiatrists. (2014). *Report of the second round of the National Audit of Schizophrenia (NAS2).* UK, Royal College of Psychiatrists.

9 Haddock G, Eisner E, Boone C, Davies G, Coogan C, Barrowclough C. (2014). An investigation of the implementation of NICE-recommended CBT interventions with people with schizophrenia. *Journal of Mental Health*. 23(4), 162-165.

10 Health and social care information centre. (2015). *Mental Health and Leaning Disabilities Data Set. UK,* Health and social care information centre.

11 Van Der Gaag, M. (2014). The efficacy of CBT for severe mental illness and the challenge of dissemination in routine care. *World Psychiatry*. 13(3), 257 – 258. **DOI**: 10.1002/wps.20162.

12 Perera G, Broadbent M, Chang CK, Downs J, Dutta R, Fernandes A, Hayes R, Henderson M, Jackson R, Jewell A, Kadra G, Little R, Pritchard M, Shetty H, Tulloch A, Stewart, R. (2016). Cohort Profile of the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) Case Register: Current Status and Recent Enhancements of an Electronic Mental Health Record Derived Data Resource. *BMJ Open*. 6(3). **DOI**:10.1136/bmjopen-2015-008721.

13 Fernandes A, Cloete D, Broadbent M, Hayes RD, Chang CK, Jackson RG, Roberts A, Tsang J, Soncul M, Liebsche J, Stewart R, Callard F. (2013). Development and Evaluation of a De-Identification Procedure for a Case Register Sourced from Mental Health Electronic Records. *BMC psychiatry*. 13(71). **DOI**: 10.1186/1472-6947-13-71.

14 Stewart R, Soremekun M, Perera G, Broadbent M, Callard F, Denis M, Hotopf M, Thornicroft G, Lovestone S. (2009). The South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) case register: development and descriptive data. *BMC Psychiatry,* 9(51). **DOI**: 10.1186/1471-244X-9-51.

15 Spyns P. (1996). Natural language processing in medicine: an overview. *Methods of information in medicine* 35.4, 285-301.

16 Cunningham H. (2002). GATE, a general architecture for text engineering. *Computers and the Humanities*, 36 (2), 223-254.

17 Hayes R, Downs J, Chang CK, Jackson R, Shetty H, Broadbent M, Hotopf M, Stewart R. (2015). The Effect of Clozapine on Premature Mortality: An Assessment of Clinical Monitoring and Other Potential Confounders. *Schizophrenia bulletin*. 41 (3): 644-655.**DOI**: 10.1093/schbul/sbu120.

18 Su YP, Chang CK, Hayes RD, Perera G, Broadbent M, David To, Hotopf M, Stewart R. (2014). Mini-mental state examination as a predictor of mortality among older people referred to secondary mental healthcare. *PLoS One,* 9(9):e105312.

19 Sultana J, Chang C-K, Hayes RD, Broadbent M, Stewart R, Corbett A, Ballard C. (2014) Associations between risk of mortality and atypical antipsychotic use in vascular dementia: a clinical cohort study. *International Journal of Geriatric Psychiatry*; 29(12), 1249-1254. **DOI**: 10.1002/gps.4101.

20 Wu C-Y, Chang C-K, Robson D, Jackson R, Chen S-J, Hayes RD, Stewart R. (2013) Evaluation of Smoking Status Identification Using Electronic Health Records and Open-Text Information in a Large Mental Health Case Register. *PLoS ONE* 8(9): e74262. **DOI**: 10.1371/journal.pone.0074262.

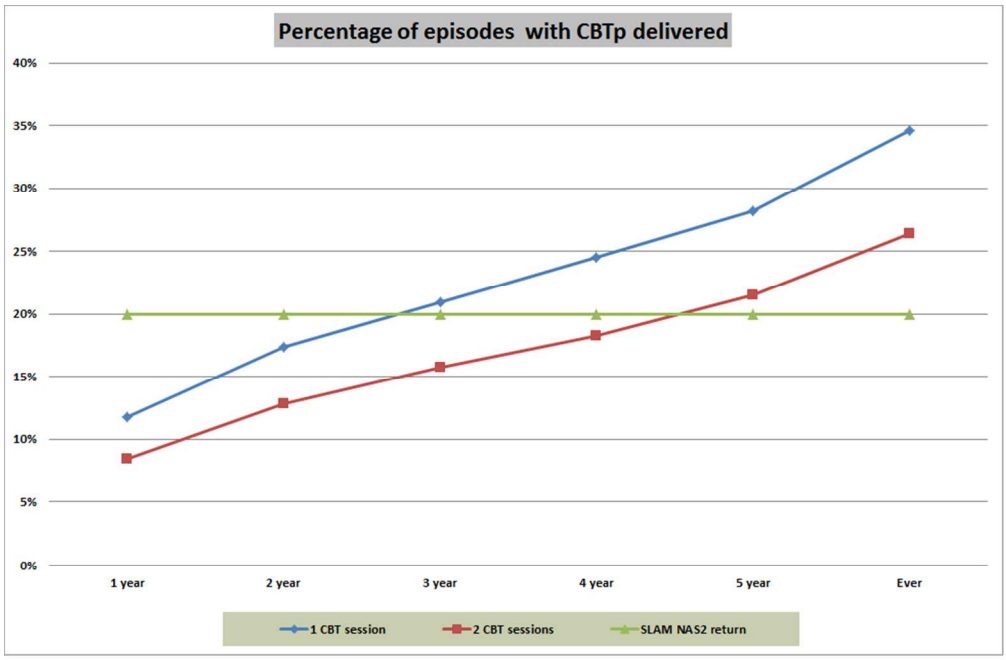

Fig 2

258x169mm (96 x 96 DPI)

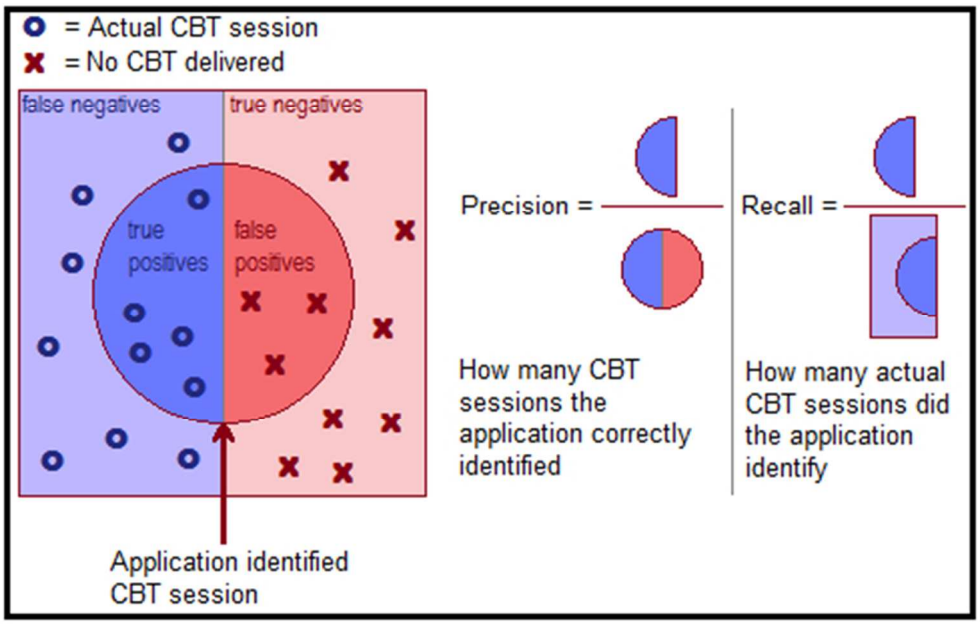

Fig 1

130x82mm (96 x 96 DPI)

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
- **Title and abstract**	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract **PAGE 2**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found **PAGE 2**
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported **PAGE 4**
Objectives	3	State specific objectives, including any prespecified hypotheses **PAGE 2**
Methods		
Study design	4	Present key elements of study design early in the paper **PAGE 6**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **PAGE 6**
Participants	6	*Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants **PAGE 9**
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic

		criteria, if applicable **PAGE 10**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group **PAGES 6,7 and 10**
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at **PAGES 9 and 10**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why **N/A**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding **PAGE 10**

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed **PAGES 12 and 13**
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders **PAGES 13 and 14**
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives **PAGE 15**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias **PAGE 16**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence **PAGE 18**
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at

http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Identification of the delivery of cognitive behavioural therapy for psychosis (CBTp) using a cross sectional sample from Electronic Health Records and Open-Text Information in a Large UK based Mental Health Case Register

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015297.R1
Article Type:	Research
Date Submitted by the Author:	24-Mar-2017
Complete List of Authors:	Colling, Craig; Kings College London (Institute of psychiatry, psychology and Neuroscience), Psychological medicine; South London and Maudsley NHS Fundation Trust Evans, Lauren; King's College London, Institute of Psychiatry; South London & Maudsley Foundation NHS Trust Broadbent, Matthew; South London and Maudsley NHS Fundation Trust; South London and Maudsley NHS Fundation Trust Chandran, David; King's College London, Institute of Psychiatry Craig, Thomas; King\'s College London (Institute of Psychiatry) kolliakou, Anna; King's College London, Institute of Psychiatry Stewart, Robert; King's College London, Institute of Psychiatry; South London and Maudsley NHS Fundation Trust Garety, Philippa; King's College London, Institute of Psychiatry
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Health informatics
Keywords:	cognitive behavioural therapy, CBT, CBT for psychosis, CBTp, electronic Health records, EHR

Identification of the delivery of cognitive behavioural therapy for psychosis (CBTp) using a cross sectional sample from Electronic Health Records and Open-Text Information in a Large UK based Mental Health Case Register

Craig Colling[1, 2], Lauren Evans[1, 2], Matthew Broadbent[1, 2], David Chandran[1], Thomas J Craig[1], Anna Kolliakou[1], Robert Stewart[1, 2] and Philippa A Garety[1, 2]

[1] Kings College London (Institute of psychiatry, psychology and Neuroscience) London, SE5 8AF, UK.

[2] South London & Maudsley Foundation NHS Trust, London, SE5 8AZ, UK.

Text (exc. figures/tables): 4,797

Author contact information – Craig Colling, Biomedical research Centre, Mapother House, Maudsley Hospital, Denmark Hill, London SE5 8AQ , craig.colling@kcl.ac.uk and (020) 3228 8584.

Abstract

Objective: Our primary objective was to identify CBT delivery for people with psychosis (CBTp) using an automated method in a large electronic health record database. We also examined what proportion of service users with a diagnosis of psychosis were recorded as having received CBTp within their episode of care provided by early intervention or promoting recovery community services for people with psychosis; during defined time periods; compared with published audits; and whether demographic characteristics differentially predicted the receipt of CBTp.

Methods: Both free text using natural language processing (NLP) techniques and structured methods of identifying CBTp were combined and evaluated for positive predictive value (PPV) and sensitivity. Using inclusion criteria from two published audits, we identified anonymised cross-sectional samples of 2,579 and 2,308 service users respectively with a case note diagnosis of schizophrenia or psychosis for further analysis.

Results: The method achieved PPV of 95% and sensitivity of 96%. Using the National Audit of Schizophrenia 2 criteria, 34.6% service users were identified as ever having received at least one session and 26.4% at least two sessions of CBTp; these are higher percentages than previously reported by manual audit of a sample for the same Trust which returned 20.0%. In the fully adjusted analysis, CBTp receipt was significantly ($p<0.05$) more likely in younger patients, in White and Other when compared with Black ethnic groups and patients with a diagnosis of Other schizophrenia spectrum and Schizoaffective disorder when compared with Schizophrenia.

Conclusions: The methods presented here provided a potential method for evaluating delivery of CBTp on a large scale, providing more scope for routine monitoring, cross-site comparisons and the promotion of equitable access.

Strengths

- Key strengths of this study were the large sample and the innovative approaches adopted to identify CBTp delivery within the clinical record

- The ability to replicate the inclusion criteria of two previous audits also allowed us to contextualise the findings, and the large data set allowed access to data by year and also to examine clinical and demographic factors influencing delivery.

- The use of routine data and automated ascertainment provide the scope for more in-depth evaluation of real-world treatment delivery and success, and the wider use of other EHR-derived data to investigate predictors of treatment receipt and outcome.

Limitations

- A limitation of this study was that it took place in a single (albeit large) service provider; however, our results have identified themes that are consistent with other findings in relation to CBTp provision

- This approach does not provide an assessment of quality of treatment

Introduction

Background

Pharmacotherapy as monotherapy for people with a diagnosis of psychosis or schizophrenia is no longer regarded as optimal treatment. The implementation of cognitive behavioural therapy for psychosis (CBTp) is of international concern and relevance [1] and CBTp, given its evidence base, is recommended in many countries including Australia and New Zealand [2], Canada [3], Spain [4] and the USA [5]. This paper is focussed on the provision of CBTp at a single UK site but the challenges associated with monitoring and improving the implementation of CBT for service users with psychosis have international relevance. For England and Wales, the NICE national guideline, recommends that psychological therapies, in particular CBTp and family intervention, should be offered; NICE makes the recommendation they are offered to all people with the diagnosis of schizophrenia and their carers[6]. However, repeatedly, within the UK, service users, charities such as Rethink,[7] policy makers and audits [8] and [9] have reported that only a small proportion of people are accessing these treatments. For example the Schizophrenia Commission reported that only about 10% of service users access CBTp.[10] To address these concerns the Department of Health and NHS England are undertaking various initiatives, including the IAPT SMI (Improving Access to Psychological therapies for severe mental illness) [11] programme and the new Early Intervention Access and Waiting Time initiative [12] both of which aim to drive up access. However, one area of uncertainty which will limit evaluation of progress, is whether we do have accurate baseline estimates of current levels of provision. A recent national audit (NAS2) [13] taking a random sample of 100 service users with a diagnosis of schizophrenia or schizoaffective disorder in the community in each of 64 participating mental health trust or health boards in England and Wales, concluded that there are significant gaps in the availability of CBTp and family interventions. For example, this manual case note audit found that Trusts reported that on average 39% of service users had been offered CBTp and 19% of service users had taken up CBTp. However there are grounds for thinking that the NAS2 audit might be inaccurate. The audit provided no definition or criteria for psychological therapy provision, asked whether a service user had *ever* been offered or received therapy, and was based on reports by consultant psychiatrists. The audit report noted that responses probably encompassed a broader set of interventions than covered by

the NICE recommendations. In contrast a detailed manual survey of a random sample of 187 records, reported a very much lower rate of offers (6.9%) and delivery (6.4%) of CBTp,[14] employing expert reviews of reported therapy record content, within a one year period, in one large mental health Trust.

Manually conducted audits of case notes and electronic records, such as NAS2, requiring individual responses of health professionals, are a labour intensive way of establishing these data, limit the number of cases that can reasonably be investigated and are too cumbersome to use routinely as practical tools to monitor service-level implementation. The UK's national minimum data set [15] does not currently require interventions to be recorded, although this may change. Although in the South London and Maudsley NHS Foundation Trust (SLaM), the site for this analysis, a structured drop down record for psychological interventions in electronic records is available, there is concern that, as non-mandatory, it is incomplete and unreliable as a means to monitor activity.

In the current study we therefore sought to develop a method of using automated text-based searches of clinical records using natural language processing (NLP) techniques, supplemented by information from structured fields, to investigate how much this might enhance our ability to provide accurate routine automatic data reports and analysis, and thus provide an efficient method of monitoring the implementation of psychological therapy provision, overcoming the limitations of manual case note audits. The decision to focus initially on CBTp delivery instead of CBTp offer was a pragmatic one based on the perceived complexity and the resultant time required for each project.

Research Question

The primary research question of the study was whether we could identify, with sufficiently high positive predictive value (PPV) and sensitivity, CBTp delivery using free text and structured methods in a large electronic service user record database. We also examined how many and what proportion of service users according to inclusion and exclusion criteria employed in published audits, with a case note diagnosis of schizophrenia or psychosis were recorded as having received CBTp within their episode of care using the CRIS database, during defined time periods, combining NLP and structured records. We then compared these data with the results of two published audits. Finally we examined whether

demographic characteristics differentially predicted the receipt of CBTp.

Methods

Setting

SLaM is a large provider of mental healthcare, serving a catchment of around 1.3 million residents in four boroughs of South London (Croydon, Lambeth, Lewisham and Southwark). The majority of people with a diagnosis of a schizophrenia spectrum disorder are served by early intervention teams for the first three years from initial presentation and by promoting recovery teams subsequently.

Study design

Source of clinical data

The data for this study were obtained from the SLaM Biomedical Research Centre (BRC) Case Register and its Clinical Record Interactive Search (CRIS) application,[16] which accesses anonymised data from the electronic health records (EHR) of individuals who have previously received or are currently receiving mental healthcare from SLaM within a robust, service user-led governance framework.[17] At the time of writing this is over 265,000 service user records. We used CRIS to replicate the inclusion criteria for NAS2 and Haddock et al (2014) as means of comparison with these two published audits. The SLaM BRC Case Register contains structured fields, such as those coding demographic information, as well as unstructured (but de-identified) free text fields from case notes and correspondence where history, mental state examination, diagnostic formulation and management plan are primarily recorded. The CRIS data resource has been approved for secondary analysis by the Oxfordshire Research Ethics Committee,[18] and a service user-led oversight committee considers all proposed research before access to the anonymised data is permitted. The electronic health record system was implemented in SLaM services in April 2006.

Overview of methodology

The initial step was to identify the delivery of CBT across all patient groups not distinguishing by diagnostic groups or other characteristics and then subsequently, and as the specific focus of this study,

to test the performance of the application for the delivery of CBT with a sample of service users with a diagnosis of psychosis (that is, 'CBTp').

Identification of CBT delivery using CRIS

Natural language processing (NLP) techniques [19] were used to identify CBT delivery from free text fields within the BRC Case Register. The annotation strategy to identify whether a clinical record was a session of CBT was developed by three human annotators (CC, LE and MB) who also completed the initial feasibility which was signed off by an expert clinical lead (PAG). All annotations were double annotated by two human annotators, and disagreements were resolved by consensus and liaising with the clinical lead if required. Inter-annotator agreement was evaluated following each batch of annotations completed and the annotation strategy was updated according to issues raised and clarifications identified. Two annotators reviewed a training set of 300 instances in the development phase before annotating a gold standard dataset of 200 where the term "CBT" (or variants of) occurred and annotated as to whether the sentence that contained the term "CBT" was an actual session of CBT rather than a historic reference to therapy, a referral for CBT, a decision not to offer CBT, or another reference to CBT which was not a therapy session. When a positive instance of CBT delivery was identified, the following features were recorded: session number, stage of treatment, the recipient of treatment and whether the CBT was delivered individually or via a group. Once the human annotations were complete, the training set was reviewed by the NLP developer (DC) to establish the rules to determine whether the CBT text is an actual session or not. These rules were coded using General Architecture Text Engineering (GATE) software.[20] Within the development process the impact of the rules applied to the training set were measured by the PPV and sensitivity. There is an inherent trade-off between the PPV, and sensitivity (as one increase the other reduces) so there is a balance between what is more important in relation to the problem domain. We concluded that for this study an evenly weighted solution was preferred with a slight preference to PPV. When PPV is prioritised, this results in false positives being minimised which increases the confidence in the test to correctly identify the positive outcome at the expense of incorrectly classifying some positive instances as false negatives.

When all the rules were developed based on the training set, the model was tested against an independent gold standard dataset to evaluate how well the model performed on unseen data using PPV and sensitivity as the metrics of evaluation. Once the mean of the PPV and sensitivity on the gold standard were greater than 85%, the resulting application was applied against the CRIS database and we further tested whether combining the NLP output with other relevant variables such as the professional group of the clinician who entered the clinical note, whether the clinical note was classified as a psychological therapy in structured data drop down menu, or whether the positioning of the CBT reference in the clinical document could be used to improve the performance of the application.

Identification of CBTp delivery using CRIS

The output of the CBT application was generated in a sample of service users with a current diagnosis of psychosis to evaluate whether the PPV and sensitivity were of an acceptable standard or whether a specific CBTp application would need to be developed.

Within SLaM, psychological interventions can be recorded through a drop down box within the clinical record, but as a non-mandatory field the recording was considered as potentially poor. To assess the quality and use of this field a senior clinician completed an assessment of 100 documents where CBT was indicated within the drop down box, identifying whether the text associated with the document could be confirmed as a session of CBT.

Both free text and structured methods of identifying CBT were combined to create a single set of results which was used for analysis purposes. As the focus of this paper is to identify the delivery of CBT for patients with a diagnosis of psychosis, the term 'CBTp' is used from this point forward.

Participants

We used the CRIS database to generate two large participant samples in this study: one replicating the inclusion criteria and the sampling time frame employed by the NAS2 audit and a second which replicated the Haddock et al (2014) audit inclusion criteria, allowing a comparison with each publicly available study.

1. **NAS2 audit inclusion criteria**

All individuals 'active' (i.e. receiving services rather than discharged from care) for at least 12 months on 01/07/2013 aged over 18 receiving either an early intervention or a promoting recovery service, with a recorded diagnosis of schizophrenia (F20.0 – F20.9) or schizoaffective disorder (F25.0 – F25.9). The NAS2 audit requested whether CBTp was "taken up" and we examined this in two ways: service users with at least one session of CBTp and service users with at least two sessions of CBTp prior to the census date.

2. **Haddock et al (2014) audit inclusion criteria**

All individuals active between 01/07/2012 and 01/07/2013 aged over 18 receiving either an early intervention or a promoting recovery service, with a recorded diagnosis of schizophrenia spectrum diagnosis (schizophrenia, schizoaffective, schizotypal and delusional disorders (F20.0 – F29.9)). CBTp delivery was defined as at least one session of CBTp within the 12 month audit period.

In addition to the original timeframe we re-sampled the data Haddock et al inclusion criteria for a separate 12 month timeframe in 2015 to check the robustness of the findings related to health inequalities.

If patients met the inclusion criteria across multiple teams within the same service type, to avoid double counting, the episodes were merged by selecting the earliest episode start date and latest end date for those episodes and presented as a single episode of care.

Demographic and service variables

The following variables were extracted for analyses: age, diagnosis, ethnicity, gender, marital status, and service type. All data obtained were the most recent prior to the census date. Ethnicity was recorded according to categories defined by the UK Office for National Statistics and categorised for analysis purposes into three groups: black (comprising black African, black Caribbean and any other black background), other (comprising white and black African, white and Asian, white and black Caribbean, any other mixed background, Indian, Pakistani, Bangladeshi, any other Asian background, Chinese, and any other ethnic group) and white (comprising white British, white Irish and any other white background). Marital status was aggregated into two groups: single/divorced (including dissolved civil partnerships and widowed) and married/co-habiting/civil partnerships. Diagnosis is routinely recorded in clinical services using the ICD-10 classification system in drop-down fields, and was limited to schizophrenia spectrum (F20 – F29), with an additional sub-grouping applied in line with the NAS2 diagnostic categories of schizophrenia (F20.0 – F20.9), schizoaffective disorder (F25.0 – F25.9) and 'other schizophrenia spectrum' (F21, F22.0 – F22.9, F23.0 – F23.9, F24, F28 and F29). We used the largest sample (using the Haddock el al inclusion criteria) to investigate the delivery of CBTp across the following categories: age group, diagnosis, gender, ethnic group, marital status and whether the patient was in contact with either the early intervention or promoting recovery service.

Statistical Analysis

Descriptive statistics for demographic variables are reported as means and standard deviations for continuous variables (age at referral) and as frequencies and percentages for all other variables. A binary logistic regression model was used to examine the differences for proportions of cases who received CBTp and those who did not. We initially undertook an unadjusted analysis by age group, diagnosis, ethnicity, gender, marital status and service type to establish whether the receipt of CBTp differed by these demographic factors. We subsequently undertook a multivariable analysis, adjusting for potential confounders by including covariates (age, diagnosis, ethnicity, gender, marital status and service type) in the model except the variable of interest. Due to the relationship between age and service type (Early

intervention services are by definition for a younger patient group) we included the partially adjusted model which excludes service as a predictor to check whether the increased likelihood of younger people receiving CBT is still present.

Results

PPV and sensitivity of identification of CBT in case records

The developed NLP CBT delivery application was evaluated against the independent gold standard resulting in PPV and sensitivity for CBT annotations of 85% and 86% respectively. Following the development of the CBT NLP application, we concluded the PPV would be improved with a tolerable reduction in sensitivity if we applied the following post-processing rule: to exclude CBT sentences that commenced after the first 200 characters of the clinical document. This post-processing rule resulted in an improved overall performance of the application, with an increase in PPV of 12% to 97% and a reduction in sensitivity of 4% to 82%. The evaluation of the structured CBT entry alone resulted in a PPV of 89%. We then combined both methods, and a measure was adopted to establish the sensitivity of the combined method by reviewing the false negatives (FN) from the NLP app and examining whether they were identified by the structured method: of the 12 FN's identified by the NLP app, 75% (9/12) were correctly identified by the structured data with the effect of increasing the sensitivity from 82% (56/68) for the NLP app alone to 96% (65/68) for the combined method. By combining methods, we therefore achieved a PPV of 97% and a sensitivity of 96%. The NLP app resulted in identifying 26% additional service users who received CBT not recorded by the drop down box.

PPV and sensitivity of identification of CBTp in case records

We further evaluated the developed NLP CBT delivery application against a sample of service users with a diagnosis of psychosis. The performance against the independent gold standard resulted in PPV and sensitivity for CBTp annotations of 81% and 85% respectively. Applying the above mentioned post processing rule (to exclude CBTp sentences that commenced after the first 200 characters of the clinical

document) resulted in an increase in PPV of 14% to 95% and a reduction in sensitivity of 7% to 78%. The evaluation of the structured CBT entry alone resulted in a PPV of 89%. Having combined both methods, of the 10 FN's identified by the NLP app 80% (8/10) were correctly identified by the structured data, with the effect of increasing the sensitivity from 78% (36/46) for the NLP app alone to 96% (44/46) for the combined method. By combining methods, we therefore achieved a PPV of 95% and sensitivity of 96%. The NLP app resulted in identifying 21% additional service users who received CBTp not recorded by the drop down box.

Delivery of CBTp using sample based on NAS2 inclusion criteria

2,308 service users were identified in the dataset as fulfilling the NAS2 inclusion criteria. Service users had a mean age of 40.7 at referral (SD 12.1; range 18-83), 60.3% (1,392/2,308) were male, 51.9% (1,197/2,308) were of a Black ethnic origin, 90.7% (2,094/2,308) were single/divorced, 78.2% (1,806/2,308) had a diagnosis of schizophrenia and 21.8% (502/2,308) had a diagnosis of schizoaffective disorder.

The SLaM return for the actual NAS2 Audit was that 20% of the random sample of N=100 were identified as ***having ever received*** CBTp. In contrast, using the current method, 34.6% (799/2,308) were identified as having at least one session and 26.4% (610/2,308) were identified as having at least two sessions of CBTp. A breakdown of CBTp delivery by diagnostic group can be viewed in Table 1.

Table 1 – CBTp delivery by diagnostic groups using NAS2 audit criteria

Diagnostic group	n	% episodes with at least 1 CBTp session	% episodes with at least 2 CBTp sessions
Schizoaffective disorder (F25.0 - F25.9)	502	42.4%	32.9%
Schizophrenia (F20.0 - F20.9)	1,806	32.4%	24.6%
Total	**2,308**	**34.6%**	**26.4%**

We also explored the level of CBTp provision by year which can be viewed in Figure 1.

Delivery of CBTp using sample based on Haddock et al inclusion criteria

2,579 service users fulfilled the inclusion criteria within the same 12 month audit period. Service users had a mean age of 40.0 at referral (SD 12.4; range 18-83), 60.3% (1,555/2,579) were male, 50.9% (1,314/2,579) were of a Black ethnic origin, 90.5% (2,339/2,579) were single/divorced, 70.0% (1,806/2,579) had a diagnosis of schizophrenia and 19.5% (502/2,579) had a diagnosis of schizoaffective disorder. We found that 12.8% (330/2,579) received CBTp interventions within the same twelve month audit period, whereas Haddock et al reported 6.4% (12/187) in their sample.

We also examined a more recent time period: 2,597 service users fulfilled the inclusion criteria within a 12 month audit period within 2015. Service users had a mean age of 39.6 at referral (SD 12.7; range 18-85), 60.4% (1,568/2,597) were male, 52.3% (1,357/2,597) were of a Black ethnic origin and 32.1% (883/2,579) were from a White ethnic origin, 90.5% (2,351/2,597) were single/divorced, 63.4% (1,646/2,597) had a diagnosis of schizophrenia and 20.0% (519/2,597) participants had a diagnosis of schizoaffective disorder. We found that 14.8% (385/2,597) received CBTp interventions within the twelve month audit period.

We additionally investigated the proportion of participants that received CBT 'year on year', by checking to see if the participants who took part in the audit in 2015 also received CBT in the 2013 audit. This check found that **13.8%** (53/385) of the participants who received CBTp in 2015 had also received CBTp in 2013.

Demographic predictors of at least one session of CBTp

The demographic characteristics of service users who received CBTp were compared with those who did not using our largest sample of N=2,579, which employed the Haddock inclusion criteria. The receipt of CBTp was more common in younger service users, in the White compared with the Black group, in those with schizoaffective disorder group compared to those with schizophrenia, and in those receiving care from the Early Intervention for psychosis teams (EI) rather than the promoting recovery teams (PR). Table

2 provides a summary of the unadjusted and adjusted multivariable logistic regression for receipt of CBTp by age group, diagnostic group, ethnic group, gender, marital status and service type.

CBTp delivery within a mental healthcare setting

Table 2 – unadjusted and adjusted logistic regressions for predictors of at least one session of CBTp

		Unadjusted			Partially adjusted*			Fully adjusted**		
Group	**n**	**Odds ratio**	**Confidence interval**	**Significance**	**Odds ratio**	**Confidence interval**	**Significance**	**Odds ratio**	**Confidence interval**	**Significance**
Age										
Under 41	1,346	1.57	1.24 – 1.99	<.001	1.57	1.23 – 2.01	<.001	1.32	1.01 – 1.72	.043
41 and over	1,233	Reference category								
Ethnicity										
Black	1,314	Reference category								
White	908	1.34	1.04 – 1.72	.024	1.40	1.08 – 1.80	.011	1.43	1.10 – 1.85	.007
Other	357	1.35	0.96 – 1.90	.081	1.33	0.94 – 1.88	.106	1.31	0.93 – 1.86	.122
Diagnosis										
Other schizophrenia spectrum	271	2.26	1.63 – 3.14	<.001	2.02	1.45 – 2.82	<.001	1.52	1.05 – 2.20	.025
Schizoaffective disorder	502	1.53	1.15 – 2.03	.003	1.47	1.10 – 1.97	.009	1.48	1.11 – 1.98	.008
Schizophrenia	1,806	Reference category								
Gender										
Male	1,555	Reference category								
Female	1,024	1.15	0.91 – 1.46	.230	1.19	0.94 - 1.52	.155	1.20	0.94 – 1.54	.139
Marital status										
Single/divorced	2,339	Reference category								
Married/co-habiting	240	0.93	0.62 – 1.40	.729	0.90	0.60 – 1.37	.623	0.95	0.63 – 1.44	.809
Service type										
Early Intervention	327	2.49	1.87 – 3.31	<.001	N/A			1.98	1.40 – 2.81	<.001
Promoting recovery	2,252	Reference category								
***Within the partially adjusted model the results were adjusted for age, ethnic group, diagnostic group, gender and marital status**										
****Within the fully adjusted model the results were adjusted for age, ethnic group, diagnostic group, gender, marital status and service**										

We additionally explored the number and percentage of participants that received CBT by the standard NHS 16 Ethnic groups to further detail the ethnic composition and CBTp provision which can be viewed in Table 3.

Table 3 – participants by Ethnic origin and CBTp delivery using largest sample

Analysis group	NHS Ethnic groups	Participants	Participants that received CBTp
Black	Black African (N)	**16.8%** (432/2579)	**9.7%** (42/432)
	Black Caribbean (M)	**14.9%** (384/2579)	**9.9%** (38/384)
	Any other black background (P)	**19.3%** (498/2579)	**13.5%** (67/498)
	Black	**50.9%** (1314/2579)	**11.2%** (147/1314)
Other	White and black Caribbean (D)	**1.4%** (37/2579)	**18.9%** (7/37)
	White and Black African (E)	**0.5%** (12/2579)	**33.3%** (4/12)
	White and Asian (F)	**0.2%** (6/2579)	**16.7%** (1/6)
	Any other mixed background (G)	**0.7%** (19/2579)	**10.5%** (2/19)
	Indian (H)	**1.4%** (36/2579)	**11.1%** (4/36)
	Pakistani (J)	**0.8%** (21/2579)	**9.5%** (2/21)
	Bangladeshi (K)	**0.5%** (12/2579)	**8.3%** (1/12)
	Any other Asian background (L)	**2.6%** (67/2579)	**16.4%** (11/67)
	Chinese (R)	**0.7%** (18/2579)	**0.0%** (0/18)
	Any other ethnic group (S)	**5.0%** (129/2579)	**15.5%** (20/129)
	Other	**13.8%** (357/2579)	**14.6%** (52/357)
White	British (A)	**27.5%** (710/2579)	**15.4%** (109/710)
	Irish (B)	**1.6%** (41/2579)	**14.6%** (6/41)
	Any other white background (C)	**6.1%** (157/2579)	**10.2%** (16/157)
	White	**35.2%** (908/2579)	**14.4%** (131/908)
		2579	**330**

Age, ethnicity, gender and marital status had a 100% completeness rate.

Discussion

To our knowledge this is the first attempt at using NLP techniques on free text entries, supplementing structured fields, in order to identify the delivery of one type of psychological therapy in a large health record data set. This was broadly successful, in that we achieved a high level of PPV (95%) and of sensitivity (96%) which is consistent with other published CRIS NLP applications, which have measured other clinical activities or characteristics such as prescribed medication,[21] Mini-Mental State Examination score,[22] diagnosis [23] and service user characteristics such as, smoking status [24] and whether the service user lived alone. [16] The methods presented here are therefore potentially effective and efficient for examining the delivery of CBTp on a large scale where manual audits are inevitably limited in sample size for logistical reasons.

NLP applications are increasingly being used to extract information from medical records for a wide range of health related areas including but not limited to the detection of adverse drug events, falls, nosocomial infections [25, 26, 27], obesity status and obesity related diseases [28, 29] and detecting patterns in patient care and patient treatment habits [30, 31] which highlights the potential for NLP to supplement other data collection methods. NLP applications for mental health services are less prominent but there have been recent studies in the US that used NLP to determine depression outcome, and adverse drug reactions, and characterisation of diagnostic profiles [32, 33, 34].

When using this method, we identified higher levels of CBTp delivery than previously reported in the SLaM contribution to the NAS2 audit using the same sampling criteria but a very much larger sample. *Note* the published audits using NAS2 and Haddock inclusion criteria differ on timeframe, diagnosis and interpretation of CBTp delivery. We also found higher levels of CBTp delivery (about double) than that reported by Haddock et al (2014) in the same time period, albeit in a different service setting. This suggests that manual audits may result in under-reporting, presumably because of the limitations of clinician knowledge or readily accessible recording in health records, and our development is encouraging because it may result in both better quality output and much less time-intensive data collection. It is notable that the NAS2 audit enquired whether CBTp had ever been provided: the methods described

here can search by year, which is clinically more useful; the data also might suggest that clinicians in responding to such an audit are typically considering perhaps the previous two years. Furthermore, when we conducted the sampling twice for 2013 and 2015 we found some evidence of a modest increase in provision - from 12.8% to 14.8%. However our results also continue to show that CBTp delivery falls very far short of the NICE recommendations of universal access. It is a matter of additional importance and concern that there do appear to be demographic predictors, suggesting access is inequitable in terms of age, diagnosis and ethnicity. Improving access to psychological therapies can be enhanced by examining data such as these and targeting provision towards under-served groups.

Strengths

Key strengths of this study were the large sample and the innovative approaches adopted to identify CBTp delivery within the clinical record. The ability to replicate the inclusion criteria of two previous audits also allowed us to contextualise the findings, and the large data set allowed access to data by year and also to examine clinical or demographic factors influencing delivery. Clearly there are also a large number of other variables in the EHR which are also potentially available for examination without the need to repeat data extraction, as would be the case in a manual audit. These might include service user characteristics, service delivery settings, therapist characteristics and aspects of therapy provision such as assessments, number of sessions, discontinuation and drop out, and clinical outcomes. The large sample size generated by this approach has enabled us identify previously unknown inequalities in the provision of CBTp within our own Trust which we have taken steps to address, such as raising with the senior team and the provision of regular monitoring reports split by demographic variables shared with clinical teams.

Limitations

A limitation of this study was that it took place in a single (albeit large) service provider; however, our results have identified themes that are consistent with other findings in relation to CBTp provision and could indicate generalisabity but would warrant further investigation. The sample presented here is reflective of the local service provision, although SLAM services may benefit from some research

funded clinical activity, the extent of which may differ to other services within the UK and internationally. However other countries such as Australia and New Zealand [2], Canada [3], Spain [4], UK and USA [5] have policies which recommend CBTp provision and therefore monitoring implementation of these policies is be of international importance. If other services were interested in adopting the method described here to identify CBTp, we would recommend that a full *de novo* evaluation of the application performance as it cannot be assumed that performance on one cohort would be directly generalisable to others [16]. A further limitation concerns the use of routine clinical data rather than *de novo* data collection. Clearly the information available is limited by what is recorded in the source records. For fully electronic health records, such as are now used routinely in UK mental health services, there are no other information repositories which provide administrative or medico-legal back-up, and therefore there are incentives for clinicians to record details of interventions, in order to provide evidence that these did actually take place. We believe that we were able to identify relevant CBT treatment receipt through the search approach used, because of the incentive for clinicians to record this and because of the limited options in the way this could be recorded which were identified through querying both structured and text fields – indeed, demonstrating that additional querying of text fields identified significantly larger numbers of episodes. However, we are not at this stage able to automate the identification of more subtle and nuanced descriptions of the treatment and its context – i.e. we could not identify the 'offer' rather than receipt of CBT, because of the wide range of wording used to record this, and we did not attempt to quantify the quality or nature of treatment received. It is possible that future advances in NLP may allow the automated ascertainment of these constructs, but it is possible that *de novo* data collection and/or manual case note evaluation will remain the only solutions, albeit limited in the samples that can be generated. Clearly an alternative approach would be to impose data collection on clinicians, by requiring them to complete structured assessments to delineate the process of offering, commencing and monitoring treatment. This would obviate the need for NLP approaches; this however, depends on clinicians' willingness to complete these instruments and for the approach to sustain itself over time – potentially problematic if clinicians also have to complete text fields for what may be seen as a more salient need to communicate information on sessions for their own and

colleagues' future reference, as well as for medico-legal purposes. It therefore seems likely that medical records data will remain a mixed economy of structured and text-derived information, and that audits will incorporate a mixture of large-scale, multi-site 'big data' analyses, and targeted in-depth case note review.

Next steps

The opportunity provided by employing methods shown here allows the proactive analysis of large EHR-derived data sets. In the future, a refinement could be to identify CBT delivery data by using data from NLP and structured fields to identify a course of CBT treatment. Initial definitions regarding the development of a course of treatment would require at least two CBT sessions with less than a three month break between sessions and in addition utilising other NLP features such as the CBT session number and stage of therapy to enhance the creation of such a construct. We are also now working on developing an application that identifies the delivery of other therapy types and applications which more precisely characterise the pathway from CBT being considered, through its offer and to actual receipt.

Contributors

PG and RS conceived the study and manuscript. CC and PG initially contributed significant text to the study manuscript. Analyses were carried out by CC, DC, LE and MB. All authors reviewed, contributed to and approved the final manuscript.

Competing interests

No competing interests.

Funding

This work was supported by the Clinical Records Interactive Search (CRIS) system funded and developed by the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London and a joint infrastructure grant from Guy's and St Thomas' Charity and the Maudsley Charity (grant number

BRC-2011-10035). PG and TJC acknowledge BRC support. CC and LE are funded by SLAM. All other authors receive salary support from the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The above funding bodies had no role in the study design; in the collection, analysis and interpretation of the data; in the writing of the report; and in the decision to submit the paper for publication.

Data sharing statement

No additional data are available.

References

1 Van Der Gaag, M. (2014). The efficacy of CBT for severe mental illness and the challenge of dissemination in routine care. *World Psychiatry*. 13(3), 257 – 258. **DOI**: 10.1002/wps.20162.

2 Galletly C, Castle D, Dark F, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the management of schizophrenia and related disorders. Aust N Z J Psychiatry. 2016:5; 410-72.

3 Canadian Psychiatric Association. Clinical practice guidelines; Treatment of Schizophrenia. 2005. https://ww1.cpa-apc.org/Publications/Clinical_Guidelines/schizophrenia/november2005/cjp-cpg-suppl1-05_full_spread.pdf (accessed March 14, 2017).

4 Working Group of the Clinical Practice Guideline for Schizophrenia and Incipient Psychotic Disorder. Mental Health Forum, coordination. Clinical Practice Guideline for Schizophrenia and Incipient Psychotic Disorder. Madrid: Quality Plan for the National Health System of the Ministry of Health and Consumer Affairs. Agency for Health Technology Assessment and Research, 2009. Clinical Practice Guideline: CAHTA. Number 2006/05-2. http://www.guiasalud.es/GPC/GPC_495_Schizophrenia_compl_en.pdf (accessed March 14, 2017).

5 Kreyenbuhl J, Buchanan RW, Dickerson FB et al. The Schizophrenia Patient Outcomes Research Team (PORT): Updated Treatment Recommendations 2009. Schizophr Bull.2010:36(1).

6 National institute clinical excellence. (2014). *Psychosis and schizophrenia in adults: prevention and management.* UK, National Collaborating Centre for Mental Health.

7 Rethink. (2008). *Your treatment, your choice survey: final report*. UK, Rethink.

8 Prytys M, Garety PA, Jolley S, Onwumere J, Craig T. (2011). Implementing the NICE guideline for schizophrenia recommendations for psychological therapies: A qualitative analysis of the attitudes of CMHT staff. *Clin Psychol Psychothera*, 18, 48–59. DOI: 10.1002/cpp.691.

9 Krupnik Y, Pilling S, Killaspy H, Dalton J. (2005). A study of family contact with clients and staff of community mental health teams. *The Psychiatrist*, 29(5), 174-176. DOI: 10.1192/pb.29.5.174

10 Schizophrenia commission. (2012). *An abandoned illness*. UK, Schizophrenia commission.

11 Department of Health. (2011). *Talking therapies: A four-year plan of action*. UK, COI.

12 Department of Health. (2014). *Achieving Better Access to Mental Health Services by 2020*. UK, COI.

13 Royal College of Psychiatrists. (2014). *Report of the second round of the National Audit of Schizophrenia (NAS2).* UK, Royal College of Psychiatrists.

14 Haddock G, Eisner E, Boone C, Davies G, Coogan C, Barrowclough C. (2014). An investigation of the implementation of NICE-recommended CBT interventions with people with schizophrenia. *Journal of Mental Health*. 23(4), 162-165.

15 Health and social care information centre. (2015). *Mental Health and Leaning Disabilities Data Set. UK,* Health and social care information centre.

16 Perera G, Broadbent M, Chang CK, Downs J, Dutta R, Fernandes A, Hayes R, Henderson M, Jackson R, Jewell A, Kadra G, Little R, Pritchard M, Shetty H, Tulloch A, Stewart, R. (2016). Cohort Profile of the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) Case Register: Current Status and Recent Enhancements of an Electronic Mental Health Record Derived Data Resource. *BMJ Open*. 6(3). **DOI**:10.1136/bmjopen-2015-008721.

17 Fernandes A, Cloete D, Broadbent M, Hayes RD, Chang CK, Jackson RG, Roberts A, Tsang J, Soncul M, Liebsche J, Stewart R, Callard F. (2013). Development and Evaluation of a De-Identification Procedure for a Case Register Sourced from Mental Health Electronic Records. *BMC psychiatry*. 13(71). **DOI**: 10.1186/1472-6947-13-71.

18 Stewart R, Soremekun M, Perera G, Broadbent M, Callard F, Denis M, Hotopf M, Thornicroft G, Lovestone S. (2009). The South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) case register: development and descriptive data. *BMC Psychiatry,* 9(51). **DOI**: 10.1186/1471-244X-9-51.

19 Spyns P. (1996). Natural language processing in medicine: an overview. *Methods of information in medicine* 35.4, 285-301.

20 Cunningham H. (2002). GATE, a general architecture for text engineering. *Computers and the Humanities*, 36 (2), 223-254.

21 Hayes R, Downs J, Chang CK, Jackson R, Shetty H, Broadbent M, Hotopf M, Stewart R. (2015). The Effect of Clozapine on Premature Mortality: An Assessment of Clinical Monitoring and Other Potential Confounders. *Schizophrenia bulletin*. 41 (3): 644-655.**DOI**: 10.1093/schbul/sbu120.

22 Su YP, Chang CK, Hayes RD, Perera G, Broadbent M, David To, Hotopf M, Stewart R. (2014). Mini-mental state examination as a predictor of mortality among older people referred to secondary mental healthcare. *PLoS One,* 9(9):e105312.

23 Sultana J, Chang C-K, Hayes RD, Broadbent M, Stewart R, Corbett A, Ballard C. (2014) Associations between risk of mortality and atypical antipsychotic use in vascular dementia: a clinical cohort study. *International Journal of Geriatric Psychiatry*; 29(12), 1249-1254. **DOI**: 10.1002/gps.4101.

24 Wu C-Y, Chang C-K, Robson D, Jackson R, Chen S-J, Hayes RD, Stewart R. (2013) Evaluation of Smoking Status Identification Using Electronic Health Records and Open-Text Information in a Large Mental Health Case Register. *PLoS ONE* 8(9): e74262. **DOI**: 10.1371/journal.pone.0074262.

25 Chazard E, Ficheur G, Merlin B, et al. PSIP consortium, Beuscart R. Detection of adverse drug events detection: data aggregation and data mining. Stud Health Technol Inform 2009;148:75–84.

26 Bates DW, Evans RS, Murff H, et al. Detecting adverse events using information technology. J Am Med Inform Assoc 2003;10:115–28.

27 Mendonca EA, Haas J, Shagina L, et al. Extracting information on pneumonia in infants using natural language processing of radiology reports. J Biomed Inform 2005;38:314–21.

28 Yang H, Spasic I, Keane JA, et al. A text mining approach to the prediction of disease status from clinical discharge summaries.

29 Guillen R. Identifying obesity and co-morbidities from medical records. Proceedings of the i2b2 Workshop on Challenges in Natural Language Processing for Clinical Data; 2009:868.

30 Pakhomov SV, Hanson PL, Bjornsen SS, et al. Automatic classification of foot examination findings using clinical notes and Machine learning. J Am Med Inform Assoc 2008;15:198–202.

31 Rao RB, Krishnan S, Niculescu RS. Data mining for improved cardiac care. SIGKDD Explor News 2006;8:3–10.

32 Perlis RH, Iosifescu DV, Castro VM, et al. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. Psychol Med 2012;42:41–50.

33 Sohn S, Kocher JP, Chute CG, et al. Drug side effect extraction from clinical narratives of psychiatry and psychology patients. J Am Med Inform Assoc 2011;18(Suppl 1):i144–9.

34 Roque FS, Jensen PB, Schmock H, et al. Using electronic patient records to discover disease correlations and stratify patient cohorts. PLoS Comput Biol 2013;7:e1002141.

Figure 1. The NAS2 audit requested whether CBTp was "taken up" and we examined this in two ways: service users with at least one session of CBTp which is represented by the blue line and service users with at least two sessions of CBTp prior to the census date which is represented by the red line split by year prior census date. The actual return for this Trust was also added as means of comparison which is represented by the green line.

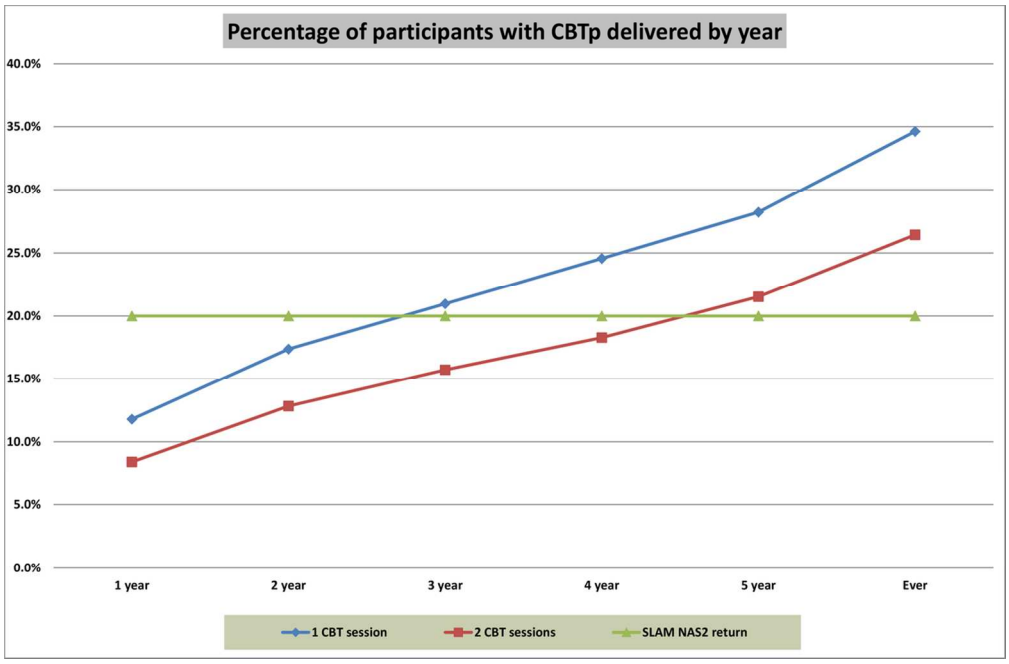

Figure 1 - CBTp delivery by year

113x74mm (300 x 300 DPI)

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
- **Title and abstract**	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract **PAGE 2**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found **PAGE 2**
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported **PAGE 4**
Objectives	3	State specific objectives, including any prespecified hypotheses **PAGE 2**
Methods		
Study design	4	Present key elements of study design early in the paper **PAGE 6**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **PAGE 6**
Participants	6	*Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants **PAGE 9**
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic

		criteria, if applicable **PAGE 10**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group **PAGES 6,7 and 10**
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at **PAGES 9 and 10**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why **N/A**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding **PAGE 10**

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed **PAGES 12 and 13**
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders **PAGES 13 and 14**
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives **PAGE 15**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias **PAGE 16**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence **PAGE 18**
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at

http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open

Identification of the delivery of cognitive behavioural therapy for psychosis (CBTp) using a cross sectional sample from Electronic Health Records and Open-Text Information in a Large UK based Mental Health Case Register

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2016-015297.R2
Article Type:	Research
Date Submitted by the Author:	05-Apr-2017
Complete List of Authors:	Colling, Craig; Kings College London (Institute of psychiatry, psychology and Neuroscience), Psychological medicine; South London and Maudsley NHS Fundation Trust Evans, Lauren; King's College London, Institute of Psychiatry; South London & Maudsley Foundation NHS Trust Broadbent, Matthew; South London and Maudsley NHS Fundation Trust; South London and Maudsley NHS Fundation Trust Chandran, David; King's College London, Institute of Psychiatry Craig, Thomas; King\'s College London (Institute of Psychiatry) kolliakou, Anna; King's College London, Institute of Psychiatry Stewart, Robert; King's College London, Institute of Psychiatry; South London and Maudsley NHS Fundation Trust Garety, Philippa; King's College London, Institute of Psychiatry
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Health informatics
Keywords:	cognitive behavioural therapy, CBT, CBT for psychosis, CBTp, electronic Health records, EHR

Identification of the delivery of cognitive behavioural therapy for psychosis (CBTp) using a cross sectional sample from Electronic Health Records and Open-Text Information in a Large UK based Mental Health Case Register

Craig Colling[1, 2], Lauren Evans[1, 2], Matthew Broadbent[1, 2], David Chandran[1], Thomas J Craig[1], Anna Kolliakou[1], Robert Stewart[1, 2] and Philippa A Garety[1, 2]

[1] Kings College London (Institute of psychiatry, psychology and Neuroscience) London, SE5 8AF, UK.

[2] South London & Maudsley Foundation NHS Trust, London, SE5 8AZ, UK.

Text (exc. figures/tables): 4,797

Author contact information – Craig Colling, Biomedical research Centre, Mapother House, Maudsley Hospital, Denmark Hill, London SE5 8AQ , craig.colling@kcl.ac.uk and (020) 3228 8584.

Abstract

Objective: Our primary objective was to identify CBT delivery for people with psychosis (CBTp) using an automated method in a large electronic health record database. We also examined what proportion of service users with a diagnosis of psychosis were recorded as having received CBTp within their episode of care provided by early intervention or promoting recovery community services for people with psychosis; during defined time periods; compared with published audits; and whether demographic characteristics differentially predicted the receipt of CBTp.

Methods: Both free text using natural language processing (NLP) techniques and structured methods of identifying CBTp were combined and evaluated for positive predictive value (PPV) and sensitivity. Using inclusion criteria from two published audits, we identified anonymised cross-sectional samples of 2,579 and 2,308 service users respectively with a case note diagnosis of schizophrenia or psychosis for further analysis.

Results: The method achieved PPV of 95% and sensitivity of 96%. Using the National Audit of Schizophrenia 2 criteria, 34.6% service users were identified as ever having received at least one session and 26.4% at least two sessions of CBTp; these are higher percentages than previously reported by manual audit of a sample for the same Trust which returned 20.0%. In the fully adjusted analysis, CBTp receipt was significantly ($p<0.05$) more likely in younger patients, in White and Other when compared with Black ethnic groups and patients with a diagnosis of Other schizophrenia spectrum and Schizoaffective disorder when compared with Schizophrenia.

Conclusions: The methods presented here provided a potential method for evaluating delivery of CBTp on a large scale, providing more scope for routine monitoring, cross-site comparisons and the promotion of equitable access.

Strengths

- Key strengths of this study were the large sample and the innovative approaches adopted to identify CBTp delivery within the clinical record

- The ability to replicate the inclusion criteria of two previous audits also allowed us to contextualise the findings, and the large data set allowed access to data by year and also to examine clinical and demographic factors influencing delivery, identifying inequalities in access which are not detectable in smaller samples.

- The use of routine data and automated ascertainment provide the scope for more in-depth evaluation of real-world treatment delivery and success, and the wider use of other EHR-derived data to investigate predictors of treatment receipt and outcome.

Limitations

- A limitation of this study was that it took place in a single (albeit large) service provider; however, our results have identified themes that are consistent with other findings in relation to CBTp provision

- This approach does not provide an assessment of quality of treatment, its specific therapeutic focus or its duration

- This approach does not identify offers of CBTp which are not taken up

Introduction

Background

Pharmacotherapy as monotherapy for people with a diagnosis of psychosis or schizophrenia is no longer regarded as optimal treatment. The implementation of cognitive behavioural therapy for psychosis (CBTp) is of international concern and relevance [1] and CBTp, given its evidence base, is recommended in many countries including Australia and New Zealand [2], Canada [3], Spain [4] and the USA [5]. This paper is focussed on the provision of CBTp at a single UK site but the challenges associated with monitoring and improving the implementation of CBT for service users with psychosis have international relevance. For England and Wales, the NICE national guideline, recommends that psychological therapies, in particular CBTp and family intervention, should be offered; NICE makes the recommendation they are offered to all people with the diagnosis of schizophrenia and their carers[6]. However, repeatedly, within the UK, service users, charities such as Rethink,[7] policy makers and audits [8] and [9] have reported that only a small proportion of people are accessing these treatments. For example the Schizophrenia Commission reported that only about 10% of service users access CBTp.[10] To address these concerns the Department of Health and NHS England are undertaking various initiatives, including the IAPT SMI (Improving Access to Psychological therapies for severe mental illness) [11] programme and the new Early Intervention Access and Waiting Time initiative [12] both of which aim to drive up access. However, one area of uncertainty which will limit evaluation of progress, is whether we do have accurate baseline estimates of current levels of provision. A recent national audit (NAS2) [13] taking a random sample of 100 service users with a diagnosis of schizophrenia or schizoaffective disorder in the community in each of 64 participating mental health trust or health boards in England and Wales, concluded that there are significant gaps in the availability of CBTp and family interventions. For example, this manual case note audit found that Trusts reported that on average 39% of service users had been offered CBTp and 19% of service users had taken up CBTp. However there are grounds for thinking that the NAS2 audit might be inaccurate. The audit provided no definition or criteria for psychological therapy provision, asked whether a service user had *ever* been offered or received therapy, and was based on reports by consultant psychiatrists. The

audit report noted that responses probably encompassed a broader set of interventions than covered by the NICE recommendations. In contrast a detailed manual survey of a random sample of 187 records, reported a very much lower rate of offers (6.9%) and delivery (6.4%) of CBTp,[14] employing expert reviews of reported therapy record content, within a one year period, in one large mental health Trust.

Manually conducted audits of case notes and electronic records, such as NAS2, requiring individual responses of health professionals, are a labour intensive way of establishing these data, limit the number of cases that can reasonably be investigated and are too cumbersome to use routinely as practical tools to monitor service-level implementation. The UK's national minimum data set [15] does not currently require interventions to be recorded, although this may change. Although in the South London and Maudsley NHS Foundation Trust (SLaM), the site for this analysis, a structured drop down record for psychological interventions in electronic records is available, there is concern that, as non-mandatory, it is incomplete and unreliable as a means to monitor activity.

In the current study we therefore sought to develop a method of using automated text-based searches of clinical records using natural language processing (NLP) techniques, supplemented by information from structured fields, to investigate how much this might enhance our ability to provide accurate routine automatic data reports and analysis, and thus provide an efficient method of monitoring the implementation of psychological therapy provision, overcoming the limitations of manual case note audits. The decision to focus initially on CBTp delivery instead of CBTp offer was a pragmatic one based on the perceived complexity and the resultant time required for each project.

Research Question

The primary research question of the study was whether we could identify, with sufficiently high positive predictive value (PPV) and sensitivity, CBTp delivery using free text and structured methods in a large electronic service user record database. We also examined how many and what proportion of service users according to inclusion and exclusion criteria employed in published audits, with a case note diagnosis of schizophrenia or psychosis were recorded as having received CBTp within their episode of care using the CRIS database, during defined time periods, combining NLP and structured records. We

then compared these data with the results of two published audits. Finally we examined whether demographic characteristics differentially predicted the receipt of CBTp.

Methods

Setting

SLaM is a large provider of mental healthcare, serving a catchment of around 1.3 million residents in four boroughs of South London (Croydon, Lambeth, Lewisham and Southwark). The majority of people with a diagnosis of a schizophrenia spectrum disorder are served by early intervention teams for the first three years from initial presentation and by promoting recovery teams subsequently.

Study design

Source of clinical data

The data for this study were obtained from the SLaM Biomedical Research Centre (BRC) Case Register and its Clinical Record Interactive Search (CRIS) application,[16] which accesses anonymised data from the electronic health records (EHR) of individuals who have previously received or are currently receiving mental healthcare from SLaM within a robust, service user-led governance framework.[17] At the time of writing this is over 265,000 service user records. We used CRIS to replicate the inclusion criteria for NAS2 and Haddock et al (2014) as means of comparison with these two published audits. The SLaM BRC Case Register contains structured fields, such as those coding demographic information, as well as unstructured (but de-identified) free text fields from case notes and correspondence where history, mental state examination, diagnostic formulation and management plan are primarily recorded. The CRIS data resource has been approved for secondary analysis by the Oxfordshire Research Ethics Committee,[18] and a service user-led oversight committee considers all proposed research before access to the anonymised data is permitted. The electronic health record system was implemented in SLaM services in April 2006.

Overview of methodology

The initial step was to identify the delivery of CBT across all patient groups not distinguishing by diagnostic groups or other characteristics and then subsequently, and as the specific focus of this study,

to test the performance of the application for the delivery of CBT with a sample of service users with a diagnosis of psychosis (that is, 'CBTp').

Identification of CBT delivery using CRIS

Natural language processing (NLP) techniques [19] were used to identify CBT delivery from free text fields within the BRC Case Register. The annotation strategy to identify whether a clinical record was a session of CBT was developed by three human annotators (CC, LE and MB) who also completed the initial feasibility which was signed off by an expert clinical lead (PAG). All annotations were double annotated by two human annotators, and disagreements were resolved by consensus and liaising with the clinical lead if required. Inter-annotator agreement was evaluated following each batch of annotations completed and the annotation strategy was updated according to issues raised and clarifications identified. Two annotators reviewed a training set of 300 instances in the development phase before annotating a gold standard dataset of 200 where the term "CBT" (or variants of) occurred and annotated as to whether the sentence that contained the term "CBT" was an actual session of CBT rather than a historic reference to therapy, a referral for CBT, a decision not to offer CBT, or another reference to CBT which was not a therapy session. When a positive instance of CBT delivery was identified, the following features were recorded: session number, stage of treatment, the recipient of treatment and whether the CBT was delivered individually or via a group. Once the human annotations were complete, the training set was reviewed by the NLP developer (DC) to establish the rules to determine whether the CBT text is an actual session or not. These rules were coded using General Architecture Text Engineering (GATE) software.[20] Within the development process the impact of the rules applied to the training set were measured by the PPV and sensitivity. There is an inherent trade-off between the PPV, and sensitivity (as one increase the other reduces) so there is a balance between what is more important in relation to the problem domain. We concluded that for this study an evenly weighted solution was preferred with a slight preference to PPV. When PPV is prioritised, this results in false positives being minimised which increases the confidence in the test to correctly identify the positive outcome at the expense of incorrectly classifying some positive instances as false negatives.

When all the rules were developed based on the training set, the model was tested against an independent gold standard dataset to evaluate how well the model performed on unseen data using PPV and sensitivity as the metrics of evaluation. Once the mean of the PPV and sensitivity on the gold standard were greater than 85%, the resulting application was applied against the CRIS database and we further tested whether combining the NLP output with other relevant variables such as the professional group of the clinician who entered the clinical note, whether the clinical note was classified as a psychological therapy in structured data drop down menu, or whether the positioning of the CBT reference in the clinical document could be used to improve the performance of the application.

Identification of CBTp delivery using CRIS

The output of the CBT application was generated in a sample of service users with a current diagnosis of psychosis to evaluate whether the PPV and sensitivity were of an acceptable standard or whether a specific CBTp application would need to be developed.

Within SLaM, psychological interventions can be recorded through a drop down box within the clinical record, but as a non-mandatory field the recording was considered as potentially poor. To assess the quality and use of this field a senior clinician completed an assessment of 100 documents where CBT was indicated within the drop down box, identifying whether the text associated with the document could be confirmed as a session of CBT.

Both free text and structured methods of identifying CBT were combined to create a single set of results which was used for analysis purposes. As the focus of this paper is to identify the delivery of CBT for patients with a diagnosis of psychosis, the term 'CBTp' is used from this point forward.

Participants

We used the CRIS database to generate two large participant samples in this study: one replicating the inclusion criteria and the sampling time frame employed by the NAS2 audit and a second which replicated the Haddock et al (2014) audit inclusion criteria, allowing a comparison with each publicly available study.

1. **NAS2 audit inclusion criteria**

All individuals 'active' (i.e. receiving services rather than discharged from care) for at least 12 months on 01/07/2013 aged over 18 receiving either an early intervention or a promoting recovery service, with a recorded diagnosis of schizophrenia (F20.0 – F20.9) or schizoaffective disorder (F25.0 – F25.9). The NAS2 audit requested whether CBTp was "taken up" and we examined this in two ways: service users with at least one session of CBTp and service users with at least two sessions of CBTp prior to the census date.

2. **Haddock et al (2014) audit inclusion criteria**

All individuals active between 01/07/2012 and 01/07/2013 aged over 18 receiving either an early intervention or a promoting recovery service, with a recorded diagnosis of schizophrenia spectrum diagnosis (schizophrenia, schizoaffective, schizotypal and delusional disorders (F20.0 – F29.9)). CBTp delivery was defined as at least one session of CBTp within the 12 month audit period.

In addition to the original timeframe we re-sampled the data Haddock et al inclusion criteria for a separate 12 month timeframe in 2015 to check the robustness of the findings related to health inequalities.

If patients met the inclusion criteria across multiple teams within the same service type, to avoid double counting, the episodes were merged by selecting the earliest episode start date and latest end date for those episodes and presented as a single episode of care.

Demographic and service variables

The following variables were extracted for analyses: age, diagnosis, ethnicity, gender, marital status, and service type. All data obtained were the most recent prior to the census date. Ethnicity was recorded according to categories defined by the UK Office for National Statistics and categorised for analysis purposes into three groups: black (comprising black African, black Caribbean and any other black background), other (comprising white and black African, white and Asian, white and black Caribbean, any other mixed background, Indian, Pakistani, Bangladeshi, any other Asian background, Chinese, and any other ethnic group) and white (comprising white British, white Irish and any other white background). Marital status was aggregated into two groups: single/divorced (including dissolved civil partnerships and widowed) and married/co-habiting/civil partnerships. Diagnosis is routinely recorded in clinical services using the ICD-10 classification system in drop-down fields, and was limited to schizophrenia spectrum (F20 – F29), with an additional sub-grouping applied in line with the NAS2 diagnostic categories of schizophrenia (F20.0 – F20.9), schizoaffective disorder (F25.0 – F25.9) and 'other schizophrenia spectrum' (F21, F22.0 – F22.9, F23.0 – F23.9, F24, F28 and F29). We used the largest sample (using the Haddock el al inclusion criteria) to investigate the delivery of CBTp across the following categories: age group, diagnosis, gender, ethnic group, marital status and whether the patient was in contact with either the early intervention or promoting recovery service.

Statistical Analysis

Descriptive statistics for demographic variables are reported as means and standard deviations for continuous variables (age at referral) and as frequencies and percentages for all other variables. A binary logistic regression model was used to examine the differences for proportions of cases who received CBTp and those who did not. We initially undertook an unadjusted analysis by age group, diagnosis, ethnicity, gender, marital status and service type to establish whether the receipt of CBTp differed by these demographic factors. We subsequently undertook a multivariable analysis, adjusting for potential confounders by including covariates (age, diagnosis, ethnicity, gender, marital status and service type) in the model except the variable of interest. Due to the relationship between age and service type (Early

intervention services are by definition for a younger patient group) we included the partially adjusted model which excludes service as a predictor to check whether the increased likelihood of younger people receiving CBT is still present.

Results

PPV and sensitivity of identification of CBT in case records

The developed NLP CBT delivery application was evaluated against the independent gold standard resulting in PPV and sensitivity for CBT annotations of 85% and 86% respectively. Following the development of the CBT NLP application, we concluded the PPV would be improved with a tolerable reduction in sensitivity if we applied the following post-processing rule: to exclude CBT sentences that commenced after the first 200 characters of the clinical document. This post-processing rule resulted in an improved overall performance of the application, with an increase in PPV of 12% to 97% and a reduction in sensitivity of 4% to 82%. The evaluation of the structured CBT entry alone resulted in a PPV of 89%. We then combined both methods, and a measure was adopted to establish the sensitivity of the combined method by reviewing the false negatives (FN) from the NLP app and examining whether they were identified by the structured method: of the 12 FN's identified by the NLP app, 75% (9/12) were correctly identified by the structured data with the effect of increasing the sensitivity from 82% (56/68) for the NLP app alone to 96% (65/68) for the combined method. By combining methods, we therefore achieved a PPV of 97% and a sensitivity of 96%. The NLP app resulted in identifying 26% additional service users who received CBT not recorded by the drop down box.

PPV and sensitivity of identification of CBTp in case records

We further evaluated the developed NLP CBT delivery application against a sample of service users with a diagnosis of psychosis. The performance against the independent gold standard resulted in PPV and sensitivity for CBTp annotations of 81% and 85% respectively. Applying the above mentioned post processing rule (to exclude CBTp sentences that commenced after the first 200 characters of the clinical

document) resulted in an increase in PPV of 14% to 95% and a reduction in sensitivity of 7% to 78%. The evaluation of the structured CBT entry alone resulted in a PPV of 89%. Having combined both methods, of the 10 FN's identified by the NLP app 80% (8/10) were correctly identified by the structured data, with the effect of increasing the sensitivity from 78% (36/46) for the NLP app alone to 96% (44/46) for the combined method. By combining methods, we therefore achieved a PPV of 95% and sensitivity of 96%. The NLP app resulted in identifying 21% additional service users who received CBTp not recorded by the drop down box.

Delivery of CBTp using sample based on NAS2 inclusion criteria

2,308 service users were identified in the dataset as fulfilling the NAS2 inclusion criteria. Service users had a mean age of 40.7 at referral (SD 12.1; range 18-83), 60.3% (1,392/2,308) were male, 51.9% (1,197/2,308) were of a Black ethnic origin, 90.7% (2,094/2,308) were single/divorced, 78.2% (1,806/2,308) had a diagnosis of schizophrenia and 21.8% (502/2,308) had a diagnosis of schizoaffective disorder.

The SLaM return for the actual NAS2 Audit was that 20% of the random sample of N=100 were identified as ***having ever received*** CBTp. In contrast, using the current method, 34.6% (799/2,308) were identified as having at least one session and 26.4% (610/2,308) were identified as having at least two sessions of CBTp. A breakdown of CBTp delivery by diagnostic group can be viewed in Table 1.

Table 1 – CBTp delivery by diagnostic groups using NAS2 audit criteria

Diagnostic group	n	% episodes with at least 1 CBTp session	% episodes with at least 2 CBTp sessions
Schizoaffective disorder (F25.0 - F25.9)	502	42.4%	32.9%
Schizophrenia (F20.0 - F20.9)	1,806	32.4%	24.6%
Total	**2,308**	**34.6%**	**26.4%**

We also explored the level of CBTp provision by year which can be viewed in Figure 1.

Delivery of CBTp using sample based on Haddock et al inclusion criteria

2,579 service users fulfilled the inclusion criteria within the same 12 month audit period. Service users had a mean age of 40.0 at referral (SD 12.4; range 18-83), 60.3% (1,555/2,579) were male, 50.9% (1,314/2,579) were of a Black ethnic origin, 90.5% (2,339/2,579) were single/divorced, 70.0% (1,806/2,579) had a diagnosis of schizophrenia and 19.5% (502/2,579) had a diagnosis of schizoaffective disorder. We found that 12.8% (330/2,579) received CBTp interventions within the same twelve month audit period, whereas Haddock et al reported 6.4% (12/187) in their sample.

We also examined a more recent time period: 2,597 service users fulfilled the inclusion criteria within a 12 month audit period within 2015. Service users had a mean age of 39.6 at referral (SD 12.7; range 18-85), 60.4% (1,568/2,597) were male, 52.3% (1,357/2,597) were of a Black ethnic origin and 32.1% (883/2,579) were from a White ethnic origin, 90.5% (2,351/2,597) were single/divorced, 63.4% (1,646/2,597) had a diagnosis of schizophrenia and 20.0% (519/2,597) participants had a diagnosis of schizoaffective disorder. We found that 14.8% (385/2,597) received CBTp interventions within the twelve month audit period.

We additionally investigated the proportion of participants that received CBT 'year on year', by checking to see if the participants who took part in the audit in 2015 also received CBT in the 2013 audit. This check found that **13.8%** (53/385) of the participants who received CBTp in 2015 had also received CBTp in 2013.

Demographic predictors of at least one session of CBTp

The demographic characteristics of service users who received CBTp were compared with those who did not using our largest sample of N=2,579, which employed the Haddock inclusion criteria. The receipt of CBTp was more common in younger service users, in the White compared with the Black group, in those with schizoaffective disorder group compared to those with schizophrenia, and in those receiving care from the Early Intervention for psychosis teams (EI) rather than the promoting recovery teams (PR). Table

2 provides a summary of the unadjusted and adjusted multivariable logistic regression for receipt of CBTp by age group, diagnostic group, ethnic group, gender, marital status and service type.

CBTp delivery within a mental healthcare setting

Table 2 – unadjusted and adjusted logistic regressions for predictors of at least one session of CBTp

		Unadjusted			Partially adjusted*			Fully adjusted**		
Group	**n**	**Odds ratio**	**Confidence interval**	**Significance**	**Odds ratio**	**Confidence interval**	**Significance**	**Odds ratio**	**Confidence interval**	**Significance**
Age										
Under 41	1,346	1.57	1.24 – 1.99	<.001	1.57	1.23 – 2.01	<.001	1.32	1.01 – 1.72	.043
41 and over	1,233	Reference category								
Ethnicity										
Black	1,314	Reference category								
White	908	1.34	1.04 – 1.72	.024	1.40	1.08 – 1.80	.011	1.43	1.10 – 1.85	.007
Other	357	1.35	0.96 – 1.90	.081	1.33	0.94 – 1.88	.106	1.31	0.93 – 1.86	.122
Diagnosis										
Other schizophrenia spectrum	271	2.26	1.63 – 3.14	<.001	2.02	1.45 – 2.82	<.001	1.52	1.05 – 2.20	.025
Schizoaffective disorder	502	1.53	1.15 – 2.03	.003	1.47	1.10 – 1.97	.009	1.48	1.11 – 1.98	.008
Schizophrenia	1,806	Reference category								
Gender										
Male	1,555	Reference category								
Female	1,024	1.15	0.91 – 1.46	.230	1.19	0.94 - 1.52	.155	1.20	0.94 – 1.54	.139
Marital status										
Single/divorced	2,339	Reference category								
Married/co-habiting	240	0.93	0.62 – 1.40	.729	0.90	0.60 – 1.37	.623	0.95	0.63 – 1.44	.809
Service type										
Early Intervention	327	2.49	1.87 – 3.31	<.001	N/A			1.98	1.40 – 2.81	<.001
Promoting recovery	2,252	Reference category								
***Within the partially adjusted model the results were adjusted for age, ethnic group, diagnostic group, gender and marital status**										
****Within the fully adjusted model the results were adjusted for age, ethnic group, diagnostic group, gender, marital status and service**										

We additionally explored the number and percentage of participants that received CBT by the standard NHS 16 Ethnic groups to further detail the ethnic composition and CBTp provision which can be viewed in Table 3.

Table 3 – participants by Ethnic origin and CBTp delivery using largest sample

Analysis group	NHS Ethnic groups	Participants	Participants that received CBTp
Black	Black African (N)	**16.8%** (432/2579)	**9.7%** (42/432)
	Black Caribbean (M)	**14.9%** (384/2579)	**9.9%** (38/384)
	Any other black background (P)	**19.3%** (498/2579)	**13.5%** (67/498)
	Black	**50.9%** (1314/2579)	**11.2%** (147/1314)
Other	White and black Caribbean (D)	**1.4%** (37/2579)	**18.9%** (7/37)
	White and Black African (E)	**0.5%** (12/2579)	**33.3%** (4/12)
	White and Asian (F)	**0.2%** (6/2579)	**16.7%** (1/6)
	Any other mixed background (G)	**0.7%** (19/2579)	**10.5%** (2/19)
	Indian (H)	**1.4%** (36/2579)	**11.1%** (4/36)
	Pakistani (J)	**0.8%** (21/2579)	**9.5%** (2/21)
	Bangladeshi (K)	**0.5%** (12/2579)	**8.3%** (1/12)
	Any other Asian background (L)	**2.6%** (67/2579)	**16.4%** (11/67)
	Chinese (R)	**0.7%** (18/2579)	**0.0%** (0/18)
	Any other ethnic group (S)	**5.0%** (129/2579)	**15.5%** (20/129)
	Other	**13.8%** (357/2579)	**14.6%** (52/357)
White	British (A)	**27.5%** (710/2579)	**15.4%** (109/710)
	Irish (B)	**1.6%** (41/2579)	**14.6%** (6/41)
	Any other white background (C)	**6.1%** (157/2579)	**10.2%** (16/157)
	White	**35.2%** (908/2579)	**14.4%** (131/908)
		2579	**330**

Age, ethnicity, gender and marital status had a 100% completeness rate.

Discussion

To our knowledge this is the first attempt at using NLP techniques on free text entries, supplementing structured fields, in order to identify the delivery of one type of psychological therapy in a large health record data set. This was broadly successful, in that we achieved a high level of PPV (95%) and of sensitivity (96%) which is consistent with other published CRIS NLP applications, which have measured other clinical activities or characteristics such as prescribed medication,[21] Mini-Mental State Examination score,[22] diagnosis [23] and service user characteristics such as, smoking status [24] and whether the service user lived alone. [16] The methods presented here are therefore potentially effective and efficient for examining the delivery of CBTp on a large scale where manual audits are inevitably limited in sample size for logistical reasons.

NLP applications are increasingly being used to extract information from medical records for a wide range of health related areas including but not limited to the detection of adverse drug events, falls, nosocomial infections [25, 26, 27], obesity status and obesity related diseases [28, 29] and detecting patterns in patient care and patient treatment habits [30, 31] which highlights the potential for NLP to supplement other data collection methods. NLP applications for mental health services are less prominent but there have been recent studies in the US that used NLP to determine depression outcome, and adverse drug reactions, and characterisation of diagnostic profiles [32, 33, 34].

When using this method, we identified higher levels of CBTp delivery than previously reported in the SLaM contribution to the NAS2 audit using the same sampling criteria but a very much larger sample. *Note* the published audits using NAS2 and Haddock inclusion criteria differ on timeframe, diagnosis and interpretation of CBTp delivery. We also found higher levels of CBTp delivery (about double) than that reported by Haddock et al (2014) in the same time period, albeit in a different service setting. This suggests that manual audits may result in under-reporting, presumably because of the limitations of clinician knowledge or readily accessible recording in health records, and our development is encouraging because it may result in both better quality output and much less time-intensive data collection. It is notable that the NAS2 audit enquired whether CBTp had ever been provided: the methods described

here can search by year, which is clinically more useful; the data also might suggest that clinicians in responding to such an audit are typically considering perhaps the previous two years. Furthermore, when we conducted the sampling twice for 2013 and 2015 we found some evidence of a modest increase in provision - from 12.8% to 14.8%. However our results also continue to show that CBTp delivery falls very far short of the NICE recommendations of universal access. It is a matter of additional importance and concern that there do appear to be demographic predictors, suggesting access is inequitable in terms of age, diagnosis and ethnicity. Improving access to psychological therapies can be enhanced by examining data such as these and targeting provision towards under-served groups. The value of informatics to monitor the delivery of psychological therapy provision and the advantages described here are important for health systems internationally.

Strengths

Key strengths of this study were the large sample and the innovative approaches adopted to identify CBTp delivery within the clinical record. The ability to replicate the inclusion criteria of two previous audits also allowed us to contextualise the findings, and the large data set allowed access to data by year and also to examine clinical or demographic factors influencing delivery. Clearly there are also a large number of other variables in the EHR which are also potentially available for examination without the need to repeat data extraction, as would be the case in a manual audit. These might include service user characteristics, service delivery settings, therapist characteristics and aspects of therapy provision such as assessments, number of sessions, discontinuation and drop out, and clinical outcomes. The large sample size generated by this approach has enabled us identify previously unknown inequalities in the provision of CBTp within our own Trust which we have taken steps to address, such as raising with the senior team and the provision of regular monitoring reports split by demographic variables shared with clinical teams.

Limitations

A limitation of this study was that it took place in a single (albeit large) service provider; however, our results have identified themes that are consistent with other findings in relation to CBTp provision and

could indicate generalisabity but would warrant further investigation. The sample presented here is reflective of the local service provision, although SLAM services may benefit from some research funded clinical activity, the extent of which may differ to other services within the UK and internationally. However other countries such as Australia and New Zealand [2], Canada [3], Spain [4], UK and USA [5] have policies which recommend CBTp provision and therefore monitoring implementation of these policies is be of international importance. If other services were interested in adopting the method described here to identify CBTp, we would recommend that a full *de novo* evaluation of the application performance as it cannot be assumed that performance on one cohort would be directly generalisable to others [16]. A further limitation concerns the use of routine clinical data rather than *de novo* data collection. Clearly the information available is limited by what is recorded in the source records. For fully electronic health records, such as are now used routinely in UK mental health services, there are no other information repositories which provide administrative or medico-legal back-up, and therefore there are incentives for clinicians to record details of interventions, in order to provide evidence that these did actually take place. We believe that we were able to identify relevant CBT treatment receipt through the search approach used, because of the incentive for clinicians to record this and because of the limited options in the way this could be recorded which were identified through querying both structured and text fields – indeed, demonstrating that additional querying of text fields identified significantly larger numbers of episodes. However, we are not at this stage able to automate the identification of more subtle and nuanced descriptions of the treatment and its context – i.e. we could not identify the 'offer' rather than receipt of CBT, because of the wide range of wording used to record this, and we did not attempt to quantify the quality or nature of treatment received. It is possible that future advances in NLP may allow the automated ascertainment of these constructs, but it is possible that *de novo* data collection and/or manual case note evaluation will remain the only solutions, albeit limited in the samples that can be generated. Clearly an alternative approach would be to impose data collection on clinicians, by requiring them to complete structured assessments to delineate the process of offering, commencing and monitoring treatment. This would obviate the need for NLP approaches; this however, depends on clinicians' willingness to complete these instruments and for the

approach to sustain itself over time – potentially problematic if clinicians also have to complete text fields for what may be seen as a more salient need to communicate information on sessions for their own and colleagues' future reference, as well as for medico-legal purposes. It therefore seems likely that medical records data will remain a mixed economy of structured and text-derived information, and that audits will incorporate a mixture of large-scale, multi-site 'big data' analyses, and targeted in-depth case note review.

Next steps

The opportunity provided by employing methods shown here allows the proactive analysis of large EHR-derived data sets. In the future, a refinement could be to identify CBT delivery data by using data from NLP and structured fields to identify a course of CBT treatment. Initial definitions regarding the development of a course of treatment would require at least two CBT sessions with less than a three month break between sessions and in addition utilising other NLP features such as the CBT session number and stage of therapy to enhance the creation of such a construct. We are also now working on developing an application that identifies the delivery of other therapy types and applications which more precisely characterise the pathway from CBT being considered, through its offer and to actual receipt.

Contributors

PG and RS conceived the study and manuscript. CC and PG initially contributed significant text to the study manuscript. Analyses were carried out by CC, DC, LE and MB. All authors reviewed, contributed to and approved the final manuscript.

Competing interests

No competing interests.

Funding

This work was supported by the Clinical Records Interactive Search (CRIS) system funded and developed by the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London and a joint infrastructure grant from Guy's and St Thomas' Charity and the Maudsley Charity (grant number BRC-2011-10035). PG and TJC acknowledge BRC support. CC and LE are funded by SLAM. All other authors receive salary support from the National Institute for Health Research (NIHR) Mental Health Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The above funding bodies had no role in the study design; in the collection, analysis and interpretation of the data; in the writing of the report; and in the decision to submit the paper for publication.

Data sharing statement

No additional data are available.

References

1 Van Der Gaag, M. (2014). The efficacy of CBT for severe mental illness and the challenge of dissemination in routine care. *World Psychiatry*. 13(3), 257 – 258. **DOI**: 10.1002/wps.20162.

2 Galletly C, Castle D, Dark F, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the management of schizophrenia and related disorders. Aust N Z J Psychiatry. 2016:5; 410-72.

3 Canadian Psychiatric Association. Clinical practice guidelines; Treatment of Schizophrenia. 2005. https://ww1.cpa-apc.org/Publications/Clinical_Guidelines/schizophrenia/november2005/cjp-cpg-suppl1-05_full_spread.pdf (accessed March 14, 2017).

4 Working Group of the Clinical Practice Guideline for Schizophrenia and Incipient Psychotic Disorder. Mental Health Forum, coordination. Clinical Practice Guideline for Schizophrenia and Incipient Psychotic Disorder. Madrid: Quality Plan for the National Health System of the Ministry of Health and Consumer Affairs. Agency for Health Technology Assessment and Research, 2009. Clinical Practice Guideline: CAHTA. Number 2006/05-2. http://www.guiasalud.es/GPC/GPC_495_Schizophrenia_compl_en.pdf (accessed March 14, 2017).

5 Kreyenbuhl J, Buchanan RW, Dickerson FB et al. The Schizophrenia Patient Outcomes Research Team (PORT): Updated Treatment Recommendations 2009. Schizophr Bull.2010:36(1).

6 National institute clinical excellence. (2014). *Psychosis and schizophrenia in adults: prevention and management.* UK, National Collaborating Centre for Mental Health.

7 Rethink. (2008). *Your treatment, your choice survey: final report*. UK, Rethink.

8 Prytys M, Garety PA, Jolley S, Onwumere J, Craig T. (2011). Implementing the NICE guideline for schizophrenia recommendations for psychological therapies: A qualitative analysis of the attitudes of CMHT staff. *Clin Psychol Psychothera*, 18, 48–59. DOI: 10.1002/cpp.691.

9 Krupnik Y, Pilling S, Killaspy H, Dalton J. (2005). A study of family contact with clients and staff of community mental health teams. *The Psychiatrist*, 29(5), 174-176. DOI: 10.1192/pb.29.5.174

10 Schizophrenia commission. (2012). *An abandoned illness*. UK, Schizophrenia commission.

11 Department of Health. (2011). *Talking therapies: A four-year plan of action*. UK, COI.

12 Department of Health. (2014). *Achieving Better Access to Mental Health Services by 2020*. UK, COI.

13 Royal College of Psychiatrists. (2014). *Report of the second round of the National Audit of Schizophrenia (NAS2).* UK, Royal College of Psychiatrists.

14 Haddock G, Eisner E, Boone C, Davies G, Coogan C, Barrowclough C. (2014). An investigation of the implementation of NICE-recommended CBT interventions with people with schizophrenia. *Journal of Mental Health*. 23(4), 162-165.

15 Health and social care information centre. (2015). *Mental Health and Leaning Disabilities Data Set. UK,* Health and social care information centre.

16 Perera G, Broadbent M, Chang CK, Downs J, Dutta R, Fernandes A, Hayes R, Henderson M, Jackson R, Jewell A, Kadra G, Little R, Pritchard M, Shetty H, Tulloch A, Stewart, R. (2016). Cohort Profile of the South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) Case Register: Current Status and Recent Enhancements of an Electronic Mental Health Record Derived Data Resource. *BMJ Open*. 6(3). **DOI**:10.1136/bmjopen-2015-008721.

17 Fernandes A, Cloete D, Broadbent M, Hayes RD, Chang CK, Jackson RG, Roberts A, Tsang J, Soncul M, Liebsche J, Stewart R, Callard F. (2013). Development and Evaluation of a De-Identification Procedure for a Case Register Sourced from Mental Health Electronic Records. *BMC psychiatry*. 13(71). **DOI**: 10.1186/1472-6947-13-71.

18 Stewart R, Soremekun M, Perera G, Broadbent M, Callard F, Denis M, Hotopf M, Thornicroft G, Lovestone S. (2009). The South London and Maudsley NHS Foundation Trust Biomedical Research Centre (SLaM BRC) case register: development and descriptive data. *BMC Psychiatry,* 9(51). **DOI**: 10.1186/1471-244X-9-51.

19 Spyns P. (1996). Natural language processing in medicine: an overview. *Methods of information in medicine* 35.4, 285-301.

20 Cunningham H. (2002). GATE, a general architecture for text engineering. *Computers and the Humanities*, 36 (2), 223-254.

21 Hayes R, Downs J, Chang CK, Jackson R, Shetty H, Broadbent M, Hotopf M, Stewart R. (2015). The Effect of Clozapine on Premature Mortality: An Assessment of Clinical Monitoring and Other Potential Confounders. *Schizophrenia bulletin*. 41 (3): 644-655.**DOI**: 10.1093/schbul/sbu120.

22 Su YP, Chang CK, Hayes RD, Perera G, Broadbent M, David To, Hotopf M, Stewart R. (2014). Mini-mental state examination as a predictor of mortality among older people referred to secondary mental healthcare. *PLoS One,* 9(9):e105312.

23 Sultana J, Chang C-K, Hayes RD, Broadbent M, Stewart R, Corbett A, Ballard C. (2014) Associations between risk of mortality and atypical antipsychotic use in vascular dementia: a clinical cohort study. *International Journal of Geriatric Psychiatry*; 29(12), 1249-1254. **DOI**: 10.1002/gps.4101.

24 Wu C-Y, Chang C-K, Robson D, Jackson R, Chen S-J, Hayes RD, Stewart R. (2013) Evaluation of Smoking Status Identification Using Electronic Health Records and Open-Text Information in a Large Mental Health Case Register. *PLoS ONE* 8(9): e74262. **DOI**: 10.1371/journal.pone.0074262.

25 Chazard E, Ficheur G, Merlin B, et al. PSIP consortium, Beuscart R. Detection of adverse drug events detection: data aggregation and data mining. Stud Health Technol Inform 2009;148:75–84.

26 Bates DW, Evans RS, Murff H, et al. Detecting adverse events using information technology. J Am Med Inform Assoc 2003;10:115–28.

27 Mendonca EA, Haas J, Shagina L, et al. Extracting information on pneumonia in infants using natural language processing of radiology reports. J Biomed Inform 2005;38:314–21.

28 Yang H, Spasic I, Keane JA, et al. A text mining approach to the prediction of disease status from clinical discharge summaries.

29 Guillen R. Identifying obesity and co-morbidities from medical records. Proceedings of the i2b2 Workshop on Challenges in Natural Language Processing for Clinical Data; 2009:868.

30 Pakhomov SV, Hanson PL, Bjornsen SS, et al. Automatic classification of foot examination findings using clinical notes and Machine learning. J Am Med Inform Assoc 2008;15:198–202.

31 Rao RB, Krishnan S, Niculescu RS. Data mining for improved cardiac care. SIGKDD Explor News 2006;8:3–10.

32 Perlis RH, Iosifescu DV, Castro VM, et al. Using electronic medical records to enable large-scale studies in psychiatry: treatment resistant depression as a model. Psychol Med 2012;42:41–50.

33 Sohn S, Kocher JP, Chute CG, et al. Drug side effect extraction from clinical narratives of psychiatry and psychology patients. J Am Med Inform Assoc 2011;18(Suppl 1):i144–9.

34 Roque FS, Jensen PB, Schmock H, et al. Using electronic patient records to discover disease correlations and stratify patient cohorts. PLoS Comput Biol 2013;7:e1002141.

Figure 1. The NAS2 audit requested whether CBTp was "taken up" and we examined this in two ways: service users with at least one session of CBTp which is represented by the blue line and service users with at least two sessions of CBTp prior to the census date which is represented by the red line split by year prior census date. The actual return for this Trust was also added as means of comparison which is represented by the green line.

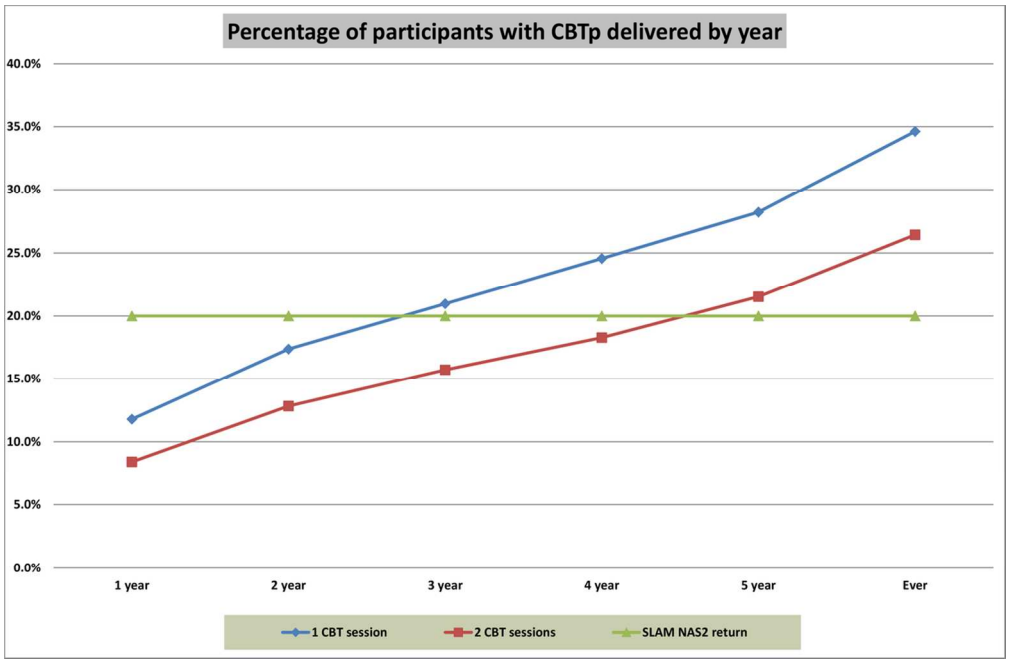

Figure 1 - CBTp delivery by year

113x74mm (300 x 300 DPI)

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation
- **Title and abstract**	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract **PAGE 2**
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found **PAGE 2**
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported **PAGE 4**
Objectives	3	State specific objectives, including any prespecified hypotheses **PAGE 2**
Methods		
Study design	4	Present key elements of study design early in the paper **PAGE 6**
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection **PAGE 6**
Participants	6	*Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants **PAGE 9**
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic

		criteria, if applicable **PAGE 10**
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group **PAGES 6,7 and 10**
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at **PAGES 9 and 10**
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why **N/A**
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding **PAGE 10**

Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed **PAGES 12 and 13**
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders **PAGES 13 and 14**
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time
		Case-control study—Report numbers in each exposure category, or summary measures of exposure
		Cross-sectional study—Report numbers of outcome events or summary measures
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives **PAGE 15**
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias **PAGE 16**
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence **PAGE 18**
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at

http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.