BMC Psychiatry

Clinical judgement, case complexity and symptom scores as predictors of outcome in depression: an exploratory analysis

M. Smith[1*], B. Francq[2], A. McConnachie[3], K. Wetherall[3], A. Pelosi[4] and J. Morrison[5]

Abstract

Background: Clinical guidelines for depression in adults recommend the use of outcome measures and stepped care models in routine care. Such measures are based on symptom severity, but response to treatment is likely to also be influenced by personal and contextual factors. This observational study of a routine clinical sample sought to examine the extent to which "symptom severity measures" and "complexity measures" assess different aspects of patient experience, and how they might relate to clinical outcomes, including disengagement from treatment.

Methods: Subjects with symptoms of depression (with or without comorbid anxiety) were recruited from people referred to an established Primary Care Mental Health Team using a stepped care model. Each participant completed three baseline symptom measures (the Personal Health Questionnaire (PHQ), Generalised Anxiety Disorder questionnaire (GAD) and Clinical Outcomes in Routine Evaluation (CORE-10)), and two assessments of "case complexity" (the Minnesota-Edinburgh Complexity Assessment Measure (MECAM) and a local complexity assessment). Clinician perception of likely completion of treatment and patient recovery was also assessed. Outcome measures were drop out and clinical improvement on the PHQ.

Results: 298 subjects were recruited to the study, of whom 258 had a sufficient dataset available for analysis. Data showed that the three measures of symptom severity used in this study (PHQ, GAD and CORE-10) seemed to be measuring distinct characteristics from those associated with the measures of case complexity (MECAM, previous and current problem count). Higher symptom severity scores were correlated with improved outcomes at the end of treatment, but there was no association between outcome and complexity measures. Clinicians could predict participant drop-out from care with some accuracy, but had no ability to predict outcome from treatment.

Conclusions: These results highlight the extent to which drop-out complicates recovery from depression with or without anxiety in real-world settings, and the need to consider other factors beyond symptom severity in planning care. The findings are discussed in relation to a growing body of literature investigating prognostic indicators in the context of models of collaborative care for depression.

Keywords: Depression, Anxiety, Recovery, Primary care, Outcomes, Drop-out

* Correspondence: Michael.Smith2@ggc.scot.nhs.uk
[1]NHS Greater Glasgow and Clyde, Glasgow, UK
Full list of author information is available at the end of the article

Background

Clinical guidelines for the management of depression recommend the use of clinical outcome measures and stepped care models in routine care [1], but it can be difficult for health services to match individual patients to the most appropriate intervention, and outcomes are often poor. Large studies in the UK and the USA have found that as few as 22% of patients receive adequate care [2], that half of patients show no response to treatment [3]; and that only 30% of cases conclude with a "planned ending" [4].

"Collaborative care" is an approach to the delivery of evidence-based mental health care for primary care patients which seeks to address these problems [5]. It includes three core elements: team-based multidisciplinary care delivery, implementation of a stepped care model (in which the intensity of care is stepped up or down depending on response to treatment), and the systematic collection of clinical outcome data to inform decisions about treatment [6–8]. The data chosen to comprise such "measurement-based care" conventionally measures depressive symptoms, adherence to treatment and side effects [9].

Guidance on depression care issued by the UK National Centre for Health and Care Excellence (NICE) recognises that "a wide range of biological, psychological and social factors, which are not captured well by current diagnostic systems, have a significant impact on the course of depression and the response to treatment" [10].

Researchers have called for investigation of treatment factors that extend beyond the choice of therapies implemented [3]. These influences include service-side factors, such as the attitude and aptitude of the treating clinicians [11]. Adherence to treatment is improved by good communication, management of expectations, patient activation and shared decision making [12–14]. Unplanned drop-out from care is an important adverse outcome, since the opportunity to modify treatment to better meet patient needs has been lost. Although important, none of these contextual influences are captured in the measures of symptom severity conventionally used to deliver collaborative care [15].

In practice, clinical assessments are based on case formulations that consider social, developmental and psychological factors, and take place in teams which should have the ability to reflect on their own style of engagement with patients.

Two forms of assessment were investigated: symptom severity measures (using the PHQ, GAD and CORE-10) and case complexity measures (using the MECAM and a bespoke case complexity instrument). The measures are described in more detail in the Methods section. This study aimed to investigate the feasibility of measuring a broader range of contextual factors when planning care for people with depression (with or without comorbid anxiety), and to assess the association of those factors with clinical outcomes, including drop out from treatment.

Methods

The service described in this paper has implemented a stepped care programme of treatment for depression in keeping with NICE guidance since its inception in 2004. The service model is described in detail elsewhere [16]. It includes the use of routine symptom severity measures (PHQ and GAD) to guide the use of brief psychological therapies by mental health clinicians working in a Primary Care Mental Health Team, but supervised from secondary care. Patients can access guided self-help, antidepressant medicines and/or (typically) 4–6 sessions of therapies such as Interpersonal Therapy and Cognitive Behavioural Therapy. Patients not responding to these interventions have prompt access to psychology, psychotherapy and/or psychiatry assessments, and ultimately to longer-term pharmacological and psychological therapies in secondary care.

Subjects with symptoms of depression defined as a score of five or more on the PHQ were recruited from people referred to a Primary Care Mental Health Team based in a town near a large urban centre in Scotland. They were included in the study whether or not they had comorbid anxiety symptoms as measured by the GAD. The PCMHT provides a service to patients aged 18 years or more, who are referred by their GP with a new case of depression, anxiety, low mood, adjustment disorder or some combination of these problems. A "new" case was defined as having been well for 6 months prior to the onset of current mood problems, or having been referred within 2 months of onset of treatment in primary care of the current episode.

People presenting with a primary alcohol problem, a primary drug problem, depression as part of bipolar affective disorder, and those with a terminal illness, medical or psychiatric emergencies or current psychosis were referred for more appropriate forms of care and hence excluded from this study, in keeping with standard local NHS procedures.

Before the first appointment, an information pack was sent to every patient outlining the clinical service they would receive, and inviting them to participate in this research study, with an information sheet and copy of the consent form. At the first appointment, clinicians reviewed the clinical and research information with the patient and invited them to consent to the research programme. If the patient chose not to consent, the appointment continued with "treatment as usual" as described in existing treatment protocols. If the patient did

consent to participate, the following assessments were completed before continuing treatment as described in existing protocols.

The assessment was in four parts:

1. Collection of demographic information (age, gender, ethnicity, postcode and employment status)
2. Completion of clinical symptom severity measures at each visit: the Personal Health Questionnaire (PHQ), Generalised Anxiety Disorder questionnaire (GAD) and Clinical Outcomes in Routine Evaluation (CORE-10)
3. Completion of complexity measures at initial visit only: the Minnesota-Edinburgh Complexity Assessment Measure (MECAM) and the local complexity measure.
4. The clinician's view about whether the patient was (a) likely to attend further treatment visits, and (b) whether they felt the patient's symptoms were likely to improve with an intervention of 5 sessions or less. These "engagement and prognosis" questions were asked at the initial visit only.

The PHQ is a self-reported, nine item questionnaire for the assessment of low mood. It incorporates DSM-IV depression diagnostic criteria, and scoring is based on the frequency of symptoms during the previous 2 weeks. It can be administered repeatedly to track the clinical course during treatment. Scores greater than 4, 9, 14 and 19 or more represent mild, moderate, moderately severe and severe depression respectively [17]. A reduction in score of five points or more is generally considered to represent a clinically significant improvement [17, 18].

The GAD is a seven item self-reported questionnaire used for assessment of generalised anxiety disorder that may be completed by the clinician or patient. Like the PHQ, scoring is based on the frequency of symptoms in the previous 2 weeks. Scores greater than 4, 9 and 14 points or more represent mild, moderate or severe anxiety [19].

The 34-item CORE-OM Clinical Outcomes in Routine Evaluation tool is a generic measure of psychological distress which covers a range of presenting problems [20]. The CORE-10 is a 10-item version of the full measure, and is used as a screening tool and outcome measure when the CORE-OM is considered too long for routine use. The measure includes two questions each about anxiety and depression, and one question each about trauma, physical problems and risk to self. A further three items enquire about day to day functioning, close relationships and social relationships. The clinical cut-off score for general psychological distress is 11 [21].

The Minnesota Edinburgh Complexity Assessment Method (MECAM [22]) is a clinician-rated measure which was designed to encourage a holistic assessment of patient needs, and initially developed for use in primary care in the UK. It asks 11 questions in four domains: Health and Wellbeing, Social Environment, Health Literacy and Communication and Action [23]. The items are scored on a four-point scale, as shown in Appendix 2.

A local indicator of case complexity was developed that could be completed by the assessing clinician at the end of the first visit. The indicator used a range of factors derived from a review of 20 team clinical case notes and relevant literature [24–27]. Indicators of case complexity comprised information that would be routinely collected as part of a standard psychiatric assessment, and included 12 "current difficulties" (such as childcare, money or housing problems), 15 "past difficulties" (such as a history of physical, sexual or emotional abuse, or parental alcohol or drug dependence), current receipt of benefits, alcohol consumption, smoking, presence of a long-term health condition, or presence of two or more health conditions. These measures were collated into counts. The full list of questions in the case complexity measure is shown in Appendix 1.

In addition, the treating clinician was asked to state at baseline whether they thought that "the patient's problems will improve with an intervention of 5 sessions or less", and whether the clinician thought that "this patient is more than 50% likely to attend their next appointment with you?"

All clinical information was held securely on NHS systems in keeping with NHS Greater Glasgow and Clyde policies on information governance.

Data for analysis was anonymised by removing names, date of birth, and the Community Health Index number (CHI; the unique identifier for all patients in the NHS in Scotland) from all research records. The CHI number was replaced with a personal research ID number independent of CHI. The key linking CHI and research database identification numbers was held securely in keeping with standard protocols.

Baseline characteristics of the participants such as gender, age, ethnicity and health conditions, as well as baseline complexity and symptom severity measures are presented descriptively. Deprivation was measured using the Scottish Index of Multiple Deprivation, the official tool used by Scottish Government to identify areas of poverty and inequality in Scotland [28]. Associations between complexity measures and symptom severity measures at baseline were examined.

An exploratory Principal Component Analysis (PCA) was performed to assess the pattern of associations between the complexity measures and symptom severity measures at baseline. Wilcoxon-Mann-Whitney tests were used to compare complexity measures and

symptom severity measures between groups defined by the clinician's two "engagement and prognosis" questions at the first visit.

We first investigated factors associated with the likelihood of drop out from the service. Drop out was defined as failure to return for treatment visits after the initial assessment visit. Drop out was summarised in relation to baseline patient characteristics, complexity and symptom severity measures, and the clinician assessment of whether the patient was likely to disengage, and associations were assessed with Fisher's Exact Test *p*-values. Univariate and multivariable logistic regression models were also used to explore these associations.

We then investigated the factors associated with treatment response, defined as achieving at least a 5 point improvement in PHQ score, amongst those who attended at least one treatment visit. This is compatible with the "reliable change index" for the PHQ used by other researchers [3]. Descriptive statistics, and Fisher's Test p-values are reported. Univariate and multivariable logistic regression models were also performed.

Results

Baseline characteristics

927 referrals were made to the service during the study period, and 576 patients attended at least one appointment, making them eligible to join the study. Of that number, 298 subjects were recruited to the study, of whom 258 had an adequate dataset available for analysis.

Participant characteristics at baseline are summarised in Table 1. Participants were 159 women and 99 men, with a mean age of 38.6 years. Participants were drawn

from each of the five deprivation quintiles in the Scottish Index of Multiple Deprivation. The distribution in this study was broadly similar to that of the catchment area. Thirty-two percent lived in the most deprived quintile, 15% in the second most deprived, 19% in the third most deprived, 10% in the fourth most deprived and 25% in the least deprived quintile. Our study population was slightly less deprived than the population in the overall catchment area of the study with 26, 15, 20, 15 and 23% in the most to least deprived quintiles respectively [29]. Nineteen percent were receiving one or more of the following benefits: job-seeker's allowance, employment support allowance, disability living allowance, or free school meals for children.

Ten percent of subjects reported that they drank to an "increased-risk" or "higher-risk" level as defined by NICE guidance (at that time, above 21 units of alcohol per week for men and 14 units per week for women) [30]. 24% of subjects reported that they currently smoked cigarettes. Sixty-four percent of participants described themselves as "healthy", with 22% reporting at least one long-term health condition and 14% having two or more long-term conditions. Just over half of participants said that they were both prescribed, and taking, antidepressant medicine.

None of the differences between male and female patients were statistically significant.

Association between baseline symptoms severity measures and case "complexity"

There were moderate, but highly significant associations between the complexity measures and symptom severity measures at baseline (Table 2).

Table 1 Baseline characteristics

		Male	Female	All
N		99	159	258
Age (years)	Mean (SD)	39.5 (13.4)	38.0 (12.4)	38.6 (12.8)
Current smoker	N (%)	30 (30%)	31 (19%)	61 (24%)
Current Antidepressant Use	N (%)	59 (60%)	81 (51%)	140 (54%)
Complexity Measures				
Past Difficulties Count	Mean (SD)	3.0 (2.5)	3.0 (2.2)	3.0 (2.3)
Current Difficulties Count	Mean (SD)	4 (2.1)	4.4 (2)	4.2 (2)
MECAM	Mean (SD)	18.9 (5.7)	19.1 (5)	19.0 (5.3)
Symptom Scores				
PHQ	Mean (SD)	16 .0 (5.6)	15.9 (5.6)	16.0 (5.6)
GAD	Mean (SD)	13.5 (4.8)	13.8 (4.7)	13.7 (4.7)
CORE-10	Mean (SD)	21.2 (6)	21.5 (6.4)	21.3 (6.2)
Clinician Assessments				
Do you think this patient is more than 50% likely to attend their next appointment with you?				
	"Yes" N (%)	86 (87%)	147 (92%)	233 (90%)
Do you think this patient's problems will improve with an intervention of 5 sessions or less?				
	"Yes" N (%)	71 (72%)	127 (80%)	198 (77%)

There were statistically signifiant associations between each of the symptom severity measures (PHQ-GAD (0.65), PHQ-CORE (0.72), GAD-CORE (0.72)) and each of the complexity measures (MECAM-Past (0.38), MECAM-Current (0.53), Past-Current (0.23)), all $p <$ 0.001).

An exploratory Principal Components Analysis of complexity measures and symptom severity measures at baseline identified two components that explained 49.7 and 20.3% of the total variation, respectively; Fig. 1 suggests that the first principal component measures general case severity, being positively correlated with all six scores. The second component differentiates between the complexity measures (positively correlated) and symptom severity measures (negatively correlated). This interpretation of the second principal component, coupled with the observation that the complexity and outcomes measures show a clear separation in terms of their patterns of correlations with the two principal components, indicate that the two groups of measures are describing distinct features of patients' clinical condition at baseline.

Complexity scores were also correlated with clinician views on prognosis at first assessment. Patients whose clinicians thought they were likely to improve with an intervention of 5 sessions or less had significantly lower problem counts and MECAM scores, and lower PHQ and CORE-10 scores (Table 3). Patients who the clinician thought were likely to continue with treatment had a lower number of current problems (though not previous problems) and lower MECAM scores. There were no significant associations between clinician assessments about probable attendance and any of the symptom severity scores at baseline.

Association between baseline measures and drop out from treatment

Being a current smoker was associated with a significantly increased likelihood of drop out from treatment, but age, gender and antidepressant use were not (Table 4). Neither the symptom severity measures nor the complexity measures were associated with dropping out from treatment. The clinician assessment about who would drop out of care was significantly associated with subsequent drop out. Although clinicians only predicted 25 subjects to be unlikely to attend the next visit, 80% of their predictions were correct. Logistic regression analyses (Table 5) broadly supported these findings.

Association between baseline measures and clinical improvement

There was a high attrition rate over time, with 159 of 258 participants returning for a second appointment, and only 39 completing the 5th treatment visit (Table 6).

Clinical improvement was defined as the achievement of a 5-point or greater reduction in PHQ score, at the last visit attended after the initial assessment, up to the 5th treatment visit, therefore the 99 participants who did not attend at least one treatment visit after the initial assessment were not included in the analyses looking at clinical improvement. Of the 159 participants that attended at least one treatment visit after the initial assessment, 98 achieved at least a 5-point improvement in PHQ score (Table 4).

Younger patients who engaged with the service were more likely to show an improvement in PHQ scores, but there was no difference by gender, or by antidepressant use (Table 4). None of the complexity measures showed any association with treatment response, but there was an association with baseline PHQ and GAD, with more severely affected patients showing the greatest improvements. Interestingly, clinicians were unable to predict which patients were most likely to improve within 5 treatment sessions. On multivariable regression, only baseline PHQ remained a significant predictor of treatment response, amongst the baseline symptom severity measures. Otherwise, only younger age remained as an independent predictor of improvement in PHQ scores.

Discussion

This exploratory study sought to investigate the associations between symptom severity scores and a range of contextual factors on clinical outcomes, including drop out from treatment. It was based on observation of a routine clinical sample, conducted in a team providing care in a collaborative, stepped care model. The patient group were a typical UK cohort in terms of demographics and symptom severity.

Table 2 Spearman correlation coefficients between baseline complexity measures and symptom severity measures, with *p*-values

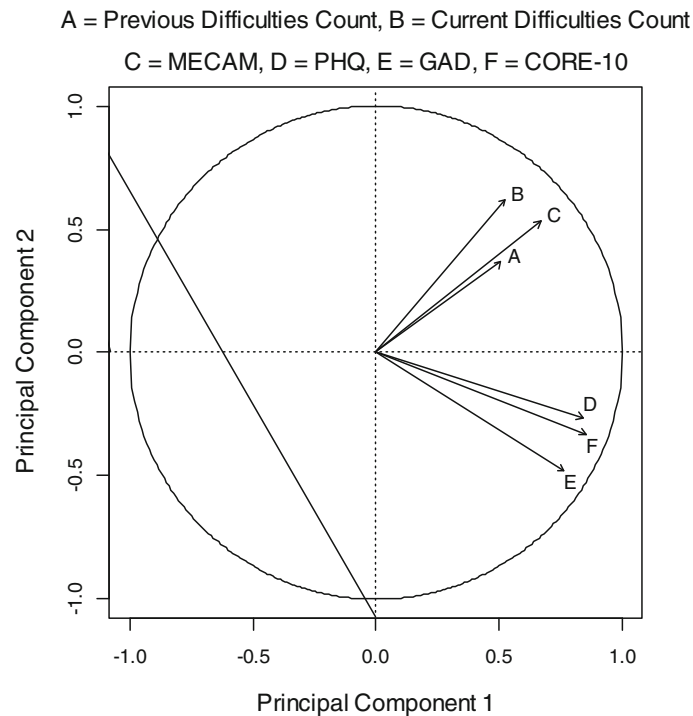|		Complexity Measures		
		Past Difficulties	Current Difficulties	MECAM
Symptom severity measures	PHQ	r = 0.24 p < 0.001	r = 0.26 p < 0.001	r = 0.41 p < 0.001
	GAD	r = 0.17 p = 0.005	r = 0.15 p = 0.016	r = 0.25 p < 0.001
	CORE-10	r = 0.28 p < 0.001	r = 0.22 p < 0.001	r = 0.35 p < 0.001

Fig. 1 Correlations between input variables and the first two components from a Principal Components Analysis of complexity measures (Previous Problem Count, Current Problem Count, MECAM) and symptom severity measures (PHQ, GAD, CORE-10) at baseline

The use of measurement-based care to improve outcomes depends on three premises: firstly, that the measurements accurately assess relevant influences on treatment, secondly that they can inform meaningful choices between treatment options, and thirdly that they can facilitate the early identification of potential harm [9]. Such harm might include clinical deterioration, emerging risks to safety or the likelihood of dropping out of treatment altogether.

Main findings

The majority of patients presented with "moderately severe" or "severe" depression, which is consistent with other studies of "low intensity" interventions [31], including the Second UK National Audit of psychological therapies [3]. Recruitment to the study was 45% of eligible subjects, comparable to other work in Improving Access to Psychological Therapies (IAPT) settings in England [32]. Drop-out rates were high, with 61% of patients attending a second appointment, and only 15% attending for five appointments. This is higher than drop-out rates of 20–40% observed in other mental health settings [33, 34], though it should be noted that the typical number of planned care visits in UK primary care mental health settings is only about six [4]. Nonetheless, the mean improvement scores on the PHQ for patients who

adhered to treatment was clinically significant and statistically significant between all treatment visits apart from between the 4th and 5th treatment visit. Symptom severity scores showed the biggest decrease between the first and second visit, with treatment continuing beyond four visits associated with less improvement.

The three measures of symptom severity (PHQ, GAD and CORE-10) were significantly correlated with each other, as were the three measures of case complexity (MECAM, previous problem count and current problem count). The symptom severity and complexity measures were also correlated with each other. However principal components analysis suggested that symptom and complexity measures at intake were associated with different aspects of patient characteristics at baseline.

Neither symptom severity nor complexity measures were associated with drop out from treatment when considered individually. On a multivariable analysis, current smoking status and the MECAM were associated with drop out. The effect with MECAM was small (OR 0.93). This is a surprising finding, since the complexity measures included socioeconomic characteristics such as unemployment or being in receipt of benefits; these would usually be considered adverse prognostic indicators. The absence of an association between experience of past adversity and drop out from treatment was

Table 3 Mean (SD) of complexity measures and symptom scores in relation to clinician assessments, with p-values from Wilcoxon-Mann-Whitney tests

	Do you think this patient's problems will improve with an intervention of 5 sessions or less?		
	Yes	No	p-value
N	198	60	
Complexity Measures			
Previous Problem Count	2.7 (2.2)	4.0 (2.5)	$p < 0.001$
Current Problem Count	4.0 (2.0)	4.8 (2.1)	$p = 0.013$
MECAM	18.3 (4.7)	21.5 (6.2)	$p < 0.001$
Symptom Scores			
PHQ	15.4 (5.6)	17.7 (5.4)	$p = 0.009$
GAD	13.5 (4.7)	14.4 (4.7)	$p = 0.161$
CORE-10	20.8 (6.2)	23.3 (5.9)	$p = 0.004$
	Do you think this patient is more than 50% likely to attend their next appointment with you?		
	Yes	No	p-value
N	233	25	
Complexity Measures			
Previous Problem Count	3.0 (2.3)	3.0 (2.2)	$p = 0.964$
Current Problem Count	4.1 (2.0)	4.8 (1.8)	$p = 0.061$
MECAM	18.7 (5.0)	21.8 (6.9)	$p = 0.040$
Symptom Scores			
PHQ	15.8 (5.5)	17.1 (6.3)	$p = 0.380$
GAD	13.6 (4.7)	14.6 (4.8)	$p = 0.270$
CORE-10	21.2 (6.3)	23.1 (5.6)	$p = 0.111$

surprising, given the extent to which Adverse Childhood Experiences (ACEs) are known to increase the prevalence of depression and the risk of poor outcomes in treatment [35, 36].

This study confirms clinician's inability to predict the likely outcome of treatment, and the importance of baseline illness severity in predicting improvement. Although the MECAM had a modest association with drop out from treatment, clinician impression was more influential. We found that clinicians could predict with some accuracy who would drop out of treatment. This effect remained influential (OR 8.59) in the multivariable analysis, although one third of patients who dropped out of treatment were not identified by the clinician. The accuracy of the clinician assessment in relation to drop out of treatment may be an early indication of difficulties in the therapeutic relationship (making drop-out more likely), or perhaps reflects a complex qualitative intuition expressed by clinicians which is not fully captured by the quantitative measures.

Baseline scores for the PHQ and GAD predicted treatment response, with more severe scores being more likely to show an improvement in PHQ of at least 5 points. On the multivariable analysis, only PHQ and younger age were associated with improvement. Clinician judgement showed no ability to determine who was most likely to improve with treatment.

Although baseline characteristics may be associated with outcome from treatment, research to date has not been able to define clear "steps" which stepped care models might use. Two recent meta-analyses using individual patient level data found that outcomes after Cognitive Behavioural Therapy (CBT) were independent of baseline severity for both CBT [37] and antidepressants or CBT [38]. The optimal treatment choices at each step may not be clear [39]. Treatment guidelines therefore differ: guidance from the National Institute for Health and Care Excellence (NICE) in England reserves antidepressant treatment for more severe depression (PHQ score 18+) [1], whereas American guidelines for the treatment of depression in primary care advocates both antidepressant and psychotherapy for patients with a PHQ score above 10 [40].

Pragmatic trials of smaller numbers of prognostic indicators [41, 42] have confirmed the importance of baseline severity of depression and anxiety, as well as other factors such as age, gender, living alone and marital status.

Strengths and limitations

One strength of the study was that it took place in a routine care setting, since participants in randomised controlled trials for depression and anxiety may not be representative of those who attend primary and psychiatric care clinics [39].

Table 4 Association between baseline patient characteristics, complexity and symptom severity measures, and clinician assessments, and patient drop out and, for those who engaged, treatment response. Drop out defined as failure to attend treatment visits after the initial assessment. N total is everyone who attended the initial visit, and N engaged is the participants that attended the initial visit and at least one treatment visit. Treatment response defined as achievement of at least a 5 point improvement in PHQ score at the last attended visit up to the 5th treatment visit. P-values from Fisher's Exact Tests

		Drop out			PHQ Response		
		N Total	N Dropout (%)	p-value	N Engaged	N Improvement (%)	p-value
	All	258	99		159	98	
Age (years)	18–29	72	33 (46%)	$p = 0.257$	39	30 (77%)	$p = 0.026$
	30–44	104	35 (34%)		69	43 (61%)	
	45+	82	31 (38%)		51	25 (47%)	
Gender	Male	99	45 (46%)	$p = 0.067$	54	35 (65%)	$p = 0.608$
	Female	159	54 (34%)		105	63 (60%)	
Current smoker	Yes	61	33 (54%)	$p = 0.006$	28	22 (79%)	$p = 0.054$
	No	197	66 (34%)		131	76 (58%)	
Current AD Use	Yes	140	55 (39%)	$p = 0.798$	85	56 (66%)	$p = 0.256$
	No	118	44 (37%)		74	42 (57%)	
Prev Prob Count	Low (≤2)	129	49 (38%)	$p = 1.000$	80	47 (59%)	$p = 0.515$
	High (≥3)	129	50 (39%)		79	51 (65%)	
Curr Prob Count	Low (≤4)	149	60 (40%)	$p = 0.518$	89	56 (63%)	$p = 0.744$
	High (≥5)	109	39 (39%)		70	42 (60%)	
MECAM	Low (≤18)	125	49 (39%)	$p = 0.799$	76	43 (57%)	$p = 0.254$
	High (≥19)	133	50 (38%)		83	55 (66%)	
PHQ	None/Mild/Moderate	98	41 (42%)	$p = 0.574$	57	26 (46%)	$p = 0.005$
	Moderate severe	79	27 (34%)		52	34 (65%)	
	Severe	125	31 (38%)		50	38 (76%)	
GAD	Minimum/Mild	54	23 (43%)	$p = 0.446$	31	15 (48%)	$p < 0.001$
	Moderate	79	33 (42%)		46	19 (41%)	
	Severe	125	43 (34%)		82	64 (78%)	
CORE-10	Low/Mild/Moderate	97	43 (44%)	$p = 0.054$	54	29 (54%)	$p = 0.096$
	Moderate severe	59	15 (25%)		44	25 (57%)	
	Severe	102	41 (40%)		61	44 (72%)	
Clinician assessment[a]	Yes	233	79 (34%)	$p < 0.001$			
	No	25	20 (80%)				
Clinician assessment[b]	Yes				129	80 (62%)	$p = 0.838$
	No				30	18 (60%)	

Table 5 Association between baseline patient characteristics, symptom scores, complexity scores, and clinician's assessment[a, b], and treatment response, defined as achievement of a 5 point reduction in PHQ-9 score at last attended visit up to the 5th treatment visit, and drop out form the service, defined as failure to return for treatment after the initial assessment visit. Data analysed using univariate and multivariable logistic regression models. Results presented as odds ratios, with 95% confidence intervals and p-values. Best-fitting multivariable models found by backward selection from all predictors

Predictor	Effect	Drop out from service		Achievement of 5 point reduction in PHQ-9	
		Univariable	Multivariable	Univariate	Multivariable
		OR (95% CI), p-value	OR (95% CI), *p*-value	OR (95% CI), p-value	OR (95% CI), p-value
Age	per 10 years	1.03 (0.85, 1.26), $p = 0.756$		0.72 (0.54, 0.94), $p = 0.017$	0.71 (0.52, 0.95), $p = 0.022$
Gender	Female vs. Male	0.62 (0.37, 1.03), $p = 0.066$		0.81 (0.41, 1.6), $p = 0.555$	
Current Smoker	Yes vs. No	2.34 (1.31, 4.22), $p = 0.004$	2.49 (1.3, 4.81), $p = 0.006$	2.65 (1.06, 7.6), $p = 0.048$	
Current AD Use	Yes vs. No	1.09 (0.66, 1.81), $p = 0.742$		1.47 (0.77, 2.81), $p = 0.239$	
Prev Prob Count	per problem	1.02 (0.92, 1.14), $p = 0.701$		1.13 (0.97, 1.32), $p = 0.131$	
Curr Prob Count	per problem	0.94 (0.82, 1.06), $p = 0.299$		1.02 (0.87, 1.2), $p = 0.82$	
MECAM	per point	0.97 (0.93, 1.02), $p = 0.275$	0.93 (0.88, 0.98), $p = 0.01$	1.03 (0.97, 1.1), $p = 0.333$	
PHQ-9	per point	1 (0.95, 1.04), $p = 0.953$		1.13 (1.07, 1.21), $p < 0.001$	1.13 (1.07, 1.21), $p < 0.001$
GAD	per point	0.96 (0.91, 1.01), $p = 0.131$		1.17 (1.09, 1.27), $p < 0.001$	
CORE-10	per point	0.98 (0.94, 1.02), $p = 0.416$		1.09 (1.04, 1.16), $p = 0.002$	
Clinician assessment[a]	No vs. Yes	7.8 (3.03, 24.13), $p < 0.001$	8.59 (3.16, 27.94), $p < 0.001$	–	
Clinician assessment[b]	Yes vs. No	–		1.09 (0.47, 2.43), $p = 0.838$	

Table 6 Mean (SD) symptom severity scores (PHQ, GAD, CORE-10) from the first 5 treatment visits (Visit 1 was the initial assessment). P-values are also included as a comparison of mean symptom severity at each visit compared with the previous visit (where applicable) using paired t-tests

	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5
N	258	159	99	63	39
PHQ	16 (5.6)	11.8 (6), $p < 0.001$	10.7 (5.9), $p < 0.001$	9.6 (5.9), $p = 0.003$	10.7 (6.3) $p = 0.369$
GAD	13.7 (4.7)	10.3 (5.4), $p < 0.001$	9.7 (5.1), $p = 0.002$	8.6 (5.3), $p < 0.001$	9.1 (5.6), $p = 0.360$
COR	21.3 (6.2)	16.8 (7.7), $p < 0.001$	15.7 (7), $p < 0.001$	13.8 (8.2), $p < 0.001$	15.2 (9.1), $p = 0.233$

However, these advantages are also associated with some weaknesses. Antidepressants and psychological therapy were used in varying combinations for different patients. The study took place in a team designed to deliver "brief interventions", which in a UK primary care mental health setting are typically of 6–10 sessions [4, 43].

Epidemiological studies suggest that the mean duration of a depressive episode is about 4 months [2]. In trials both natural remission (43% within 6 months) and placebo response rates (49%) in primary care trials are high [15, 44]; but this uncontrolled study was not able to investigate their effects.

Studies of Adverse Childhood Experiences show that not all negative life experiences have the same impact on later-life outcomes [45], yet in this study there was no adjustment for the severity or chronicity of those experiences. Likewise, no enquiry was made about potential protective factors, even though these may influence outcomes [46, 47].

Conclusions

This study highlights the importance of drop-out from care as an important therapeutic consideration. Drop-out has been found to be more likely early in treatment, particularly after the second visit. Strategies to minimise drop out include attending to health literacy and patient expectations of treatment [48, 49], establishing a therapeutic alliance [50], and of "person-centred" approaches to care and shared decision making [51].

King commented that "real-world data show that outcomes from psychotherapy are messy and difficult to predict" [52]. This observational study confirms that condition severity is the best indicator of outcome, and finds that clinician judgement may be a useful predictor of potential drop out from treatment. Further research into the characteristics of clinician prediction might usefully complement other measures used in collaborative care.

Supplementary information

> **Additional file 1.** Appendix 1: Local "Complexity" measures. Appendix 2: MECAM items.

the long-term outcome of depression. All data will be disposed of securely after the retention period.

Consent for publication

Not applicable: individual patient data is not reported in this study.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]NHS Greater Glasgow and Clyde, Glasgow, UK. [2]Institute of Statistics, Biostatistics and Actuarial Sciences, Université Catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium. [3]Robertson Centre for Biostatistics, University of Glasgow, Glasgow, UK. [4]The Priory Hospital, Glasgow, UK. [5]Senate Office, University of Glasgow, Glasgow, UK.

References

1. National Institute for Health and Clinical Excellence. Depression in adults: treatment and management. Full guideline, Version 1.0. 2017.
2. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The Epidemiology of Major Depressive Disorder: Results from the National Comorbidity Survey Replication (NCS-R). J Am Med Assoc. 2003;289:3095–105.
3. Pybis J, Saxon D, Hill A, Barkham M. The comparative effectiveness and efficiency of cognitive behaviour therapy and generic counselling in the treatment of depression: evidence from the 2nd UK National Audit of psychological therapies. BMC Psychiatry. 2017;17(1):215.
4. Stiles WB, Barkham M, Wheeler S. Effect of duration of psychological therapy on recovery and improvement rates: evidence from UK routine practice†. Br J Psychiatry. 2015;207(2):115–22.
5. Vanderlip ER, Rundell J, Avery M, Alter C, Engel C, Fortney J, et al Dissemination of integrated care within adult primary care settings. The collaborative care model. American Psychiatric Association, Academy of Pyschosomatic Medicine. 2016.
6. Ramanuj P, Ferenchick EK, Pincus HA. Depression in primary care: part 2-management. BMJ (Clinical research ed.). Br Med J Publishing Group. 2019; 365:l835.
7. Gunn J, Diggens J, Hegarty K, Blashki G. A systematic review of complex system interventions designed to increase recovery from depression in primary care. BMC Health Serv Res. 2006;6:88.
8. Bower P, Gilbody S, Richards D, Fletcher J, Sutton A. Collaborative care for depression in primary care. Making sense of a complex intervention: systematic review and meta-regression. Br J Psychiatry. 2006;189:484–93.
9. Trivedi MH. Evaluating and monitoring treatment response in depression using measurement-based assessment and rating scales. J Clin Psychiatry. 2013;74(7):e14.
10. National Institute for Health and Clinical Excellence. Depression in Adults. NICE clinical guideline 90. NHS Evidence; 2009. p. 5.
11. Okiishi J, Lambert MJ, Nielsen SL, Ogles BM. Waiting for supershrink: an empirical analysis of therapist effects. Clin Psychol Psychother. 2003;10(6): 361–73.
12. Barry MJ, Edgman-Levitan S. Shared decision making — the pinnacle of patient-centered care. N Engl J Med. 2012;366(9):780–1.
13. Lee EO, Emanuel EJ. Shared decision making to improve care and reduce costs. N Engl J Med. 2013;368(1):6–8.
14. Bell RA, Kravitz RL, Thom D, Krupat E, Azari R. Unmet expectations for care and the patient-physician relationship. J Gen Intern Med. 2002; 17(11):817–24.
15. Arroll B, Chin W, Martis W, Goodyear-Smith F, Mount V, Kingsford D, et al. Antidepressants for treatment of depression in primary care: a systematic review and meta-analysis. J Prim Health Care. 2016;8(4):325.
16. Smith MJ, Ackland L, O'Loughlin S, Young D, Pelosi AJ, Morrison J. 'Doing well': description of a complex intervention to improve depression care. Prim Health Care Res Dev. 2010:1–13.
17. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13.
18. McMillan D, Gilbody S, Richards D. Defining successful treatment outcome in depression using the PHQ-9: a comparison of methods. J Affect Disord. 2010 Dec;127(1–3):122–9.
19. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166(10): 1092–7.
20. CORE-IMS, The CORE Outcome Measure (CORE-OM) [Internet]. [cited 2020 March 09]. Available from: http://www.coreims.co.uk/About_Core_System_ Outcome_Measure.html.
21. Barkham M, Bewick B, Mullin T, Gilbody S, Connell J, Cahill J, et al. The CORE-10: a short measure of psychological distress for routine use in the psychological therapies. Couns Psychother Res. 2013;13(1):3–13.
22. Maxwell M, Hibberd C, Pratt R, Cameron I, Mercer S. Development and initial validation of the Minnesota Edinburgh complexity assessment method (MECAM) for use within the keep well health check. NHS Health Scotland; 2011.
23. Maxwell M, Cameron I, Mercer S. Development and initial validation of the Minnesota Edinburgh complexity assessment method (MECAM) for use within the keep well health check. 2011;.
24. Cochrane R, Robertson A. The life events inventory: a measure of the relative severity of psycho-social stressors. J Psychosom Res. 1973;17(2):135–9.
25. Bottomley C, Nazareth I, Torres-González F, Svab I, Maaroos H-I, Geerlings MI, et al. Comparison of risk factors for the onset and maintenance of depression. Br J Psychiatry. 2010;196(1):13–7.
26. Brugha T, Bebbington P, Tennant C, Hurry J. The list of threatening experiences: a subset of 12 life event categories with considerable long-term contextual threat. Psychol Med. 1985;15(01):189–94.
27. Hatch SL, Mishra G, Hotopf M, Jones PB, Kuh D. Appraisals of stressors and common mental disorder from early to mid-adulthood in the 1946 British birth cohort. J Affect Disord. 2009;119(1–3):66–75.
28. National Statistics Scotland. IntroducingThe Scottish Index of Multiple Deprivation 2016. Edinburgh; 2016.
29. Scottish Government. Local authority summary SIMD 2012 Renfrewshire. Edinburgh; 2012.
30. National Institute for Health and Care Excellence. Alcohol-use disorders: prevention (PH24). Manchester; 2010.
31. Bower P, Kontopantelis E, Sutton A, Kendrick T, Richards DA, Gilbody S, et al. Influence of initial severity of depression on effectiveness of low intensity interventions: meta-analysis of individual patient data. BMJ. 2013;346(feb26 2):f540.
32. Hepgul N, King S, Amarasinghe M, Breen G, Grant N, Grey N, et al. Clinical characteristics of patients assessed within an improving access to psychological therapies (IAPT) service: results from a naturalistic cohort study (predicting outcome following psychological therapy; PROMPT). BMC Psychiatry. 2016;16(1):52.
33. Wells JE, Browne MO, Aguilar-Gaxiola S, Al-Hamzawi A, Alonso J, Angermeyer MC, et al. Drop out from out-patient mental healthcare in the World Health Organization's world Menta health survey initiative. Br J Psychiatry. 2013;202(1):42–9.
34. Henzen A, Moeglin C, Giannakopoulos P, Sentissi O. Determinants of dropout in a community-based mental health crisis Centre. BMC Psychiatry. 2016;16:111.
35. De Venter M, Demyttenaere K, Bruffaerts R. The relationship between adverse childhood experiences and mental health in adulthood. A systematic literature review. Tijdschr Psychiatr. 2013;55(4):259–68.
36. Nanni V, Uher R, Danese A. Childhood maltreatment predicts unfavorable course of illness and treatment outcome in depression: a meta-analysis. Am J Psychiatry. 2012;169(2):141–51.
37. Furukawa TA, Weitz ES, Tanaka S, Hollon SD, Hofmann SG, Andersson G, et al. Initial severity of depression and efficacy of cognitive–behavioural therapy: individual-participant data meta-analysis of pill-placebo-controlled trials. Br J Psychiatry. 2017;210(3):190–6.
38. Weitz ES, Hollon SD, Twisk J, van Straten A, Huibers MJH, David D, et al. Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy. JAMA Psychiatry. 2015; 72(11):1102.
39. Wisniewski SR, Rush AJ, Nierenberg AA, Gaynes BN, Warden D, Luther JF, et al. Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR*D report. Am J Psychiatry. 2009;166(5):599–607.
40. Trangle M, Gursky J, Haight R, Hardwig J, Hinnenkamp T, Kessler D, Mack N MM. Adult depression in primary care. 2016.

41. Chondros P, Davidson S, Wolfe R, Gilchrist G, Dowrick C, Griffiths F, et al. Development of a prognostic model for predicting depression severity in adult primary patients with depressive symptoms using the diamond longitudinal study. J Affect Disord. 2018;227:854–60.
42. Angstman KB, Garrison GM, Gonzalez CA, Cozine DW, Cozine EW, Katzelnick DJ. Prediction of primary care depression outcomes at six months: validation of DOC-6 ©. J Am Board Fam Med. 2017;30(3):281–7.
43. King M. Duration of psychotherapy has little association with outcome. Br J Psychiatry. 2015;207(2).
44. Stegenga BT, Kamphuis MH, King M, Nazareth I, Geerlings MI. The natural course and outcome of major depressive disorder in primary care: the PREDICT-NL study. Soc Psychiatry Psychiatr Epidemiol. 2012;47(1):87–95.
45. Finkelhor D, Shattuck A, Turner H, Hamby S. Improving the adverse childhood experiences study scale. JAMA Pediatr. 2013;167(1):70.
46. McPherson KE, Kerr S, McGee E, Morgan A, Cheater FM, McLean J, et al. The association between social capital and mental health and behavioural problems in children and adolescents: an integrative systematic review. BMC Psychol. 2014;2(1):7.
47. Bellis MA, Hardcastle K, Ford K, Hughes K, Ashton K, Quigg Z, et al. Does continuous trusted adult support in childhood impart life-course resilience against adverse childhood experiences - a retrospective study on adult health-harming behaviours and mental well-being. BMC Psychiatry. 2017; 17(1):140.
48. Kravitz RL. Patient engagement programs for recognition and initial treatment of depression in primary care. JAMA. 2013;310(17):1818.
49. Parsons S, Winterbottom A, Cross P, Redding D. The quality of patient engagement and involvement in primary care. London: The Kings Fund; 2010.
50. Horvath AO, Del Re AC, Flückiger C, Symonds D. Alliance in individual psychotherapy. Psychotherapy. 2011;48(1):9–16.
51. Dixon LB, Holoshitz Y, Nossel I. Treatment engagement of individuals experiencing mental illness: review and update. World Psychiatry. 2016; 15(1):13–20.
52. King M. Duration of psychotherapy has little association with outcome. Br J Psychiatry. 2015;207(2):93–4.

Publisher's Note