Original Paper

The burden associated with, and management of, difficult-to-treat depression in patients under specialist psychiatric care in the United Kingdom

Journal of Psychopharmacology
2022, Vol. 36(5) 545–556

DOI: 10.1177/02698811221090628
journals.sagepub.com/home/jop

Tiago Costa[1,2], Bayar Menzat[3,4], Tomas Engelthaler[3], Benjamin Fell[3], Tarso Franarin[3], Gloria Roque[3], Yiran Wei[3], Xinyue Zhang[3] and R Hamish McAllister-Williams[1,2]

Abstract

Background: Major depressive disorder (MDD) is common and often has sub-optimal response to treatment. Difficult-to-treat depression (DTD) is a new concept that describes 'depression that continues to cause significant burden despite usual treatment efforts'.

Aims: To identify patients with likely DTD in UK secondary care and examine demographic, disease and treatment data as compared with 'non-DTD' MDD patients.

Methods: Anonymised electronic health records (EHRs) of five specialist mental health National Health Service (NHS) Trusts in the United Kingdom were analysed using a natural language processing model. Data on disease characteristics, comorbidities and treatment histories were extracted from structured fields and using natural language algorithms from unstructured fields. Patients with MDD aged ⩾18 years were included in the analysis; those with presumed DTD were identified on the basis of MDD history (duration and recurrence) and number of treatments prescribed.

Results: In a sample of 28,184 patients with MDD, 19% met criteria for DTD. Compared to the non-DTD group, patients with DTD were more likely to have severe depression, suicidal ideation, and comorbid psychiatric and/or physical illness, as well as higher rates of hospitalisation. They were also more likely to be in receipt of unemployment and sickness/disability benefits. More intensive treatment strategies were used in the DTD group, including higher rates of combination therapy, augmentation, psychotherapy and electroconvulsive therapy.

Conclusion: This study demonstrates the feasibility of identifying patients with probable DTD from EHRs and highlights the increased burden associated with MDD in these patients.

Keywords

Difficult-to-treat depression, treatment-resistant depression, antidepressants, drugs for depression, drugs for psychosis, drugs for relapse prevention, clinical management, burden of disease, burden of illness, hospital admission

Introduction

Major depressive disorder (MDD) is one of the biggest causes of disability worldwide, affecting 300 million people, the equivalent to 4.4% of the world's population (World Health Organization (WHO), 2017). In 2015, depressive disorders led to a global total of over 50 million years lived with disability (WHO, 2017). In England, the total cost of services for depression (health and social care, criminal justice services and informal care from family members) was estimated to be in the region of £1.7 billion; adding lost employment increased this by a further £5.8 billion – 2007 data from McCrone et al. (2008).

MDD is often considered as an episodic condition. If an episode is causing significant functional impairment, it may warrant treatment using psychotherapeutic, pharmacological or neurostimulatory treatments (Malhi and Mann, 2018). However, a significant proportion of patients have sub-optimal responses to treatment and experience MDD as a chronic condition. In the largest clinical trial ever conducted of what would be consistent with first-line UK treatment of MDD, the selective serotonin reuptake inhibitor (SSRI) citalopram led to around 30% of patients being relatively free of symptoms, that is, in remission (Trivedi et al., 2006). For those who do not respond, an alternative treatment can be tried, but a significant proportion of patients do not achieve remission despite serial treatments (Rush et al., 2006). Such patients are often described in the literature as suffering from 'treatment-resistant depression' (TRD). TRD is an arbitrary academic construct with no universally accepted

[1]Northern Centre for Mood Disorders, Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK
[2]Cumbria, Northumberland, Tyne and Wear NHS Foundation Trust, Newcastle upon Tyne, UK
[3]Akrivia Health, Oxford Centre for Innovation, Oxford, UK
[4]Etcembly Ltd, Magdalen Centre, Oxford, UK

Corresponding author:
R Hamish McAllister-Williams, Northern Centre for Mood Disorders, Translational and Clinical Research Institute, Newcastle University, Wolfson Research Centre, Campus for Ageing and Vitality, Newcastle upon Tyne NE4 5PL, UK.
Email: hamish.mcallister-williams@newcastle.ac.uk

criteria. It is based on failure to achieve a response or remission from a series of sequential acute treatment trials. Most commonly, TRD is defined as a failure to respond to adequate trials of two different treatments (Brown et al., 2019; Gaynes et al., 2020) but there is lack of consensus regarding the required length of trial, dose of drug for depression and outcome measures to assess response for those treatment trials. Even the 'classic' criteria of failure to respond to two drugs for depression ('antidepressants'), the most common element in TRD definitions, was present in just 50.3% of 150 trials systematically reviewed (Brown et al., 2019). TRD definitions also rarely take into account non-response to psychotherapy or neurostimulatory treatments (Brown et al., 2019; McAllister-Williams et al., 2018). In addition, the concept of TRD does not address issues of lack of tolerability of treatment or the key issue of a non-sustained response/remission, the risk of which is known to be higher the more previous failed treatment trials a patient has had (Rush et al., 2006). The first recommendation of a Delphi-method-based consensus guideline for the definition of TRD (Sforzini et al., 2021) was the need for a 'definition of TRD for clinical trials conducted for regulatory purposes', highlighting the challenge arising from the lack of clear criteria. It has been proposed that a more heuristic conceptualisation is of 'difficult-to-treat depression' (DTD) based upon a clinically determined situation in which a patient is suffering from MDD that continues to be associated with a significant burden despite usual treatment interventions (McAllister-Williams et al., 2020; Rush et al., 2019). A DTD model is more pragmatic, drawing on the models of care for chronic physical health problems with waxing and waning symptoms such as arthritis, diabetes and hypertension. DTD is not a diagnosis per se, but rather a framework or model of care (McAllister-Williams et al., 2021a) arguably more appropriate for examining the burden of depression in naturalistic clinical practice.

MDD is highly comorbid with many other mental health conditions including anxiety, post-traumatic stress disorder (PTSD), eating disorders and substance misuse (Rush et al., 2005). It is also frequently comorbid with a range of physical disorders such as type 2 diabetes, asthma, gastrointestinal and musculoskeletal conditions (Gagnon and Patten, 2002; Nouwen et al., 2010). There are strong suggestions of a particular association with cardiovascular dysfunction (Cai et al., 2019; Feng et al., 2019). Conversely, depression is twice as likely in those individuals with multimorbidity than those with only one comorbidity, and three times as likely than those with no comorbidity (Read et al., 2017). The presence of comorbidity is a major driver of the economic burden of MDD, with 62% of total health care costs being due to comorbid conditions rather than MDD itself (Greenberg et al., 2015).

These two factors of treatment failure and comorbidity interact. Data from North America (Li et al., 2019), Sweden (Reutfors et al., 2018) and Hungary (Dome et al., 2021) have shown that patients with TRD have significantly higher all-cause mortality than other depressed patients. In addition, more treatment failures, and thus a greater degree of TRD, are associated with higher health care costs (Johnston et al., 2019; Russell et al., 2004).

Observational data from real-world clinical practice, outside of the confines of controlled clinical trials, are crucial to define what constitutes 'standard' or usual treatment efforts, both for MDD and DTD. It is also relevant to identify all comorbidities (physical, psychiatric, substance misuse and iatrogenic) and respective prescribed treatments, as these will impact on the patient's trajectory. This basic understanding is a fundamental initial step for the research, design and commission of clinical services for patients with DTD (Martin-Cook et al., 2021; Rush et al., 2022).

The aim of this study was to identify a group of patients presumed to have DTD in UK specialist mental health National Health Service (NHS) Trusts and to examine demographic, disease and treatment data through the analysis of secondary-care mental health records ('real-world data'). DTD was defined on the basis of the history of patients' MDD and number of treatments prescribed. Clearly a spectrum of both degree of difficulty of treatment and burden of illness is likely in a population of patients with MDD. Any dichotomisation used to define whether patients are deemed to have DTD or not is arbitrary and as such has its limitations. Selecting for DTD based on duration of illness, number of previous episodes and treatment history will inevitably lead to differential burden. In addition, the analysis was cross-sectional and hence some patients may not have yet met our DTD criteria but will do over further time. While recognising the limitations of this abstraction, we compared the demographic, disease and treatment characteristics associated with DTD and non-DTD to explore what the particular issues associated with DTD might be.

Methods

Design and sample

An initial feasibility study was conducted using de-identified data from a single specialist mental health NHS Trust: Oxford Health NHS Foundation Trust. This allowed for an estimation of the prevalence of DTD in the study population and to establish the most useful indicators of service utilisation. Patients were then identified from five specialist mental health NHS Trusts: Oxford Health NHS Foundation Trust; Cumbria, Northumberland, Tyne and Wear NHS Foundation Trust; Southern Health NHS Foundation Trust; South West London & St George's Mental Health Trust; and West London NHS Trust. This was a convenience sample. To identify patients, a retrospective audit of de-identified electronic health records (EHRs) was performed using the UK-Clinical Records Interactive Search (UK-CRIS) system (Vaci et al., 2020).

There is no strict operational definition of DTD. Rather it is described as depression persisting despite usual treatment efforts (McAllister-Williams et al., 2020). For this analysis, a pragmatic approach has been taken to the identification of patients referred to specialist mental health services with depression that appears to be difficult-to-treat on the basis of clinical characteristics. Factors included chronicity of the current episode or frequency of episodes, and a history of multiple antidepressive treatments (psychotherapy, medication or electroconvulsive therapy (ECT)). To identify such patients, a series of filters were applied to the total patient population of the five NHS Trusts (Figure 1). The overall study population comprised those with a primary diagnosis of depression, defined as International Classification of Diseases, Tenth Revision (ICD-10) codes F32.x or F33.x, who were aged 18 years or over. Within this overall study population, patients were identified as having presumed DTD if they met all of the following criteria: the first episode of depression

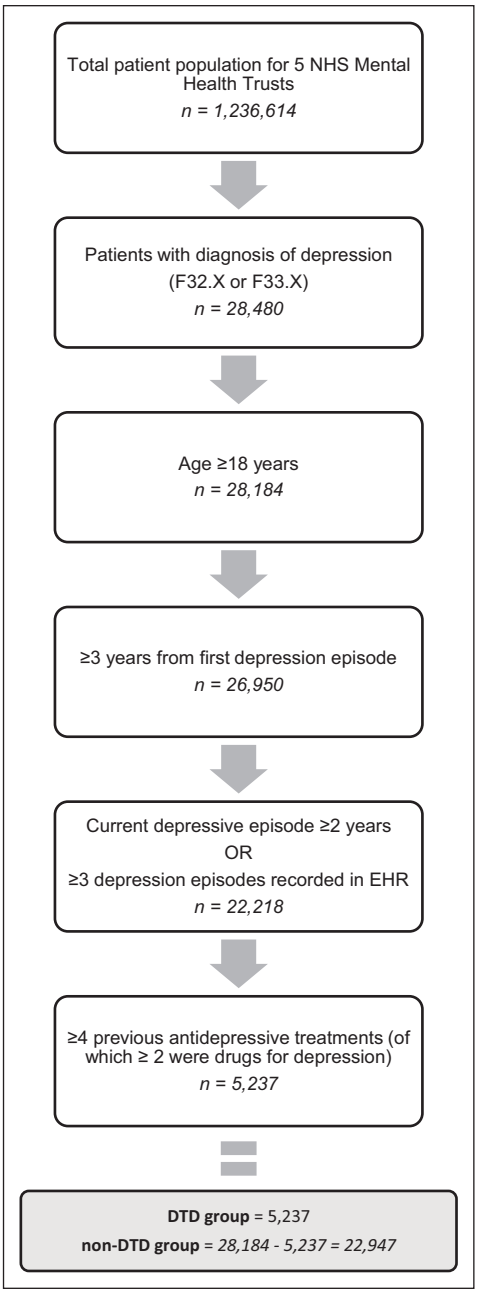

Figure 1. Patient selection flow diagram.

had been coded at least 3 years previously; the patient had a current chronic episode of depression (defined as lasting ⩾2 years) or they had had multiple recurrences (defined as ⩾3 episodes recorded in their EHR); problems with finding a tolerable and effective treatment as indexed by having had ⩾4 antidepressive treatments, of which ⩾2 were drugs for depression ('antidepressants', excluding augmentation). For example, the patient could have received two drugs for depression (either two courses of drugs for depression in monotherapy with/without augmentation or a combination regimen of drugs for depression), one course of psychotherapy and one course of ECT. The choice of a 3-year window for data collection was pragmatic. Given that the data examining burden were for the preceding 3 years, the first episode of depression had to have been coded at least 3 years previously. The use of ⩾4 antidepressive treatments as a selection threshold for the DTD group is based upon an element of the criteria proposed for the threshold for 'multi-therapy resistant MDD' (McAllister-Williams et al., 2018).

Given the chronic and relapsing nature of DTD, and the coding of clinical data in NHS secondary-care mental health services, the definition of what constitutes an episode is challenging. In the data set examined, a referral could be to a community mental health team (CMHT), crisis team, psychiatric liaison service or any other element of the relevant Trust. If a patient is transferred from one element of the service (e.g. crisis team) to another (e.g. CMHT), they are recorded as being discharged by one and referred to the other. This movement between elements of services is not necessarily temporally contiguous with episodes of illness. As a result, to identify episodes of illness rather than episodes of care in different elements of services, for this analysis, an episode was defined as a new referral to any element of secondary care ⩾6 months after the last discharge, again from any element of services.

Patients were excluded if they had a coded ICD-10 diagnosis of bipolar affective disorder (F31.*), manic episode (F30.*), any of the schizophrenic, schizotypal and delusional disorders (F2*.*; with the exception of F23.*) or dementia (F00.*, F01.*, F02.*, F03.* or G30.*). Patients were also excluded if they had no ICD-10 diagnosis coded or if they had no progress notes allowing for data extraction.

Data were extracted from the EHRs of all eligible patients through to 1 February 2021. All data elements were extracted from the entire period covered by patients' EHRs, which started from the point of referral to secondary psychiatric care, or the point at which electronic records were instigated by the Trust for patients whose referrals predated use of EHRs. All EHRs meeting the above criteria from the participating Trusts were included for analysis. Living status was not an inclusion criterion so some patients with recorded dates of death (176 DTD and 1230 non-DTD patients) were included.

The analysis was approved through the UK-CRIS application system, with each participating NHS Trust submitting the proposal to their internal governance committees for review. All analysts with access to the project data were vetted by the data controller(s) and completed the requisite information governance training according to the requirements of each NHS site.

Data and measures

Data were collected from anonymised patients' EHRs by Akrivia Health (formerly CRISTal Health; Oxford, UK) using the UK-CRIS system, which includes natural language algorithms that can mine unstructured data (clinical notes, letters, documents and other free text fields), providing a more complete and representative answer to research questions than can be extracted from structured reports, forms or other formal internal data-gathering tools (Vaci et al., 2020).

The Med7 deep learning natural language processing (NLP) model (Kormilitzin et al., 2021) was used to extract data on pharmacological treatments (psychotropics and medications prescribed for physical health). It creates structured fields such as

Table 1. Socio-demographic characteristics in DTD and non-DTD groups.

	DTD group ($n=5237$)	Non-DTD group ($n=22,947$)	p value	Effect size
Age, years				
Mean (SD)	52.8 (17.3)	51.6 (20.1)	$p<0.001$[a]	$d=0.06$[b]
Median (IQR)	53 (40, 64)	51 (35, 65)		
Gender, n (%)[c]			$p<0.001$[d]	df=1
Female	3241 (61.9)	13,557 (59.1)		$V=0.02$[e]
Male	1988 (38.0)	9337 (40.7)		
Ethnicity, n (%)			$p<0.001$[f]	df=1
White	4340 (82.9)	15,797 (68.8)		$V=0.12$[g]
Asian	287 (5.5)	1732 (7.5)		
Black	145 (2.8)	734 (3.2)		
Mixed	108 (2.1)	476 (2.1)		
Other	357 (6.8)	4208 (18.3)		
Employment, n (%)			$p<0.001$[d]	df=5
Employed	1446 (21.9)	7114 (31.0)		$V=0.15$[h]
Unemployed	1547 (29.5)	7214 (31.4)		
Retired	247 (4.7)	1292 (5.6)		
Student	332 (6.3)	2263 (9.9)		
Sickness or disability benefits	1496 (28.6)	3708 (16.2)		
Other/not known	469 (9.0)	1356 (5.9)		

df: degrees of freedom; IQR: interquartile range; SD: standard deviation.
[a]*T* test.
[b]Cohen's *d* value.
[c]Gender not known for 8 and 53 patients in DTD and non-DTD groups, respectively.
[d]Chi-square test with Bonferroni correction for multiple comparisons.
[e]Cramer's *V* statistic for proportion of female patients in DTD/non-DTD groups (at 1 df $V<0.1$ represents very small effect size).
[f]Chi-square test, with white as reference category.
[g]Cramer's *V* statistic for proportion of white patients in DTD/non-DTD groups (at 1 df $V\approx 0.1$ represents a small effect size).
[h]Cramer's *V* statistic for proportions of patients with each status in DTD/non-DTD groups (at 5 df $V=0.15$ represents medium effect size).

the medication name, the dose and the nature of the mention, that is, whether the drug was discussed in terms of current medication or in some other context. The Med7 NLP model was originally trained on general physical health data (Johnson et al., 2016), being then fine-tuned on the UK-CRIS EHR data to learn domain knowledge of mental health records (Kormilitzin et al., 2021). Only data related to current (at the time of entry into the EHRs) medication were extracted. The extracted data were then pre-filtered by cross-referencing the output against a matching table of variables of interest (in simplistic terms, a list of names of drugs, including alternative names such as brand /generic names; Supplementary Table 1), allowing also for the correction of obvious input errors such as misspellings of drug names. Not all medication records in clinical notes indicate a dosage, so in some cases it was not possible to determine the dose prescribed. There was no data imputation in the study database; the analysis only included doses which could be extracted directly from the clinical notes using Med7. Some entries contain non-plausible data due to clinician input error (e.g. a recorded dosage of 2000 mg for a medication with a recommended dosage of 50–200 mg). To address this limitation, a filter for dose outliers outside the 95% confidence interval around the mean dose for that drug was applied (i.e. records outside the 2.5% and 97.5% quantile thresholds were identified and removed).

In addition to information on pharmacological treatments abstracted from unstructured EHR text fields via Med7, the following de-identified data elements were drawn directly from structured EHR fields: socio-demographics (age, gender, ethnicity and employment status); depression characteristics (International Classification of Diseases (ICD) diagnosis, duration, recurrence, severity and psychotic features); comorbid diagnoses (mental and physical health); and hospital admissions for psychiatric conditions. Disease duration was calculated from the 'diagnosis start date' and 'diagnosis end date' fields in the EHR. Finally, dedicated rule-based pattern-matching NLP algorithms were developed for this project by Akrivia Health to extract information on employment status, history of suicidality or self-harm, referrals for psychotherapy and number of ECT referrals.

A key indicator of burden, and also a potential confounder for other data, is individual patients' duration of EHRs. EHR duration was calculated in two ways. First, EHR duration was calculated as the time elapsed between the date of the oldest recorded progress note and 1 February 2021. If the patient died before 1 February 2021, the date of death was used as the EHR end date. We excluded patients with a total EHR duration >50 years as they were outliers and this seems improbable, as well as patients with missing progress note dates. While this method maintains a consistent end date (the only exception being death), the analysis may include patients who are no longer actively receiving care from specialist services. To account for closed referrals, a second method for calculating EHR duration was used based on 'referrals tables' that detail dates of referrals to various elements of

specialist service and the date of such referrals being 'removed' or closed. This method calculated the time elapsed between the date of the first referral received and date of last referral removed. If patients had 'open' referrals by 1 February 2021, that was used as the EHR end date. Again, outliers with EHR length >50 years were excluded, as well as patients with missing or erroneous referral dates (e.g. last referral removed before first referral accepted).

Admission rates (average admissions per patient per year) were calculated as the total number of admissions during total service stay (first admission to last discharge), divided by the total service stay in years. If the service stay was <1 year, it was coded as 1 year for the purposes of calculating the rate (i.e. a patient with one admission and a total service stay of ⩽1 year will have an admission rate of 1). Admission rates were calculated both for those patients who had been admitted at least once and for the entire DTD and non-DTD groups. Patients were excluded from the calculation if the admission dates were implausible (discharge date post-dated the admission date).

Statistical analysis

Descriptive statistics were performed for all variables. Univariate inferential statistic tests were used to compare the DTD and non-DTD groups. For continuous outcome variables (age, medication dosages, frequency and duration of hospital admission), the groups were compared using independent samples *t* tests. Effect sizes were estimated based on Cohen's *d* values, with *d* values >0.5 considered to represent potentially meaningful differences. All other outcome variables were categorical; groups were compared using chi-square tests, with alpha thresholds adjusted for multiple comparisons using the Bonferroni correction. Effect sizes for differences in socio-demographic parameters, disease characteristics and use of non-pharmacological interventions were estimated using Cramer's *V* statistic, with values around 0.1, 0.3 and 0.5 indicating small, medium and large effect sizes, respectively, for categorical variables with 1 degree of freedom.

Results

Data were obtained for 28,184 patients with depression (ICD-10 F32.X or F33.X), 5237 (19%) of whom met criteria for presumed DTD, with the remaining 22,947 patients comprising the non-DTD group (Figure 1).

EHR duration

EHR duration, calculated as the time elapsed between the date of the oldest recorded progress note and 1 February 2021 (or patient's date of death if before this date, which was the case in 176 DTD and 1230 non-DTD patients), had means and medians, respectively, of 124 and 121 months for the DTD group and 101 and 104 months for the non-DTD group. We excluded 7 DTD and 7 non-DTD patients with a total EHR duration >50 years, as well as 92 non-DTD patients with missing progress note dates (none in DTD group). When EHR duration was calculated as the time elapsed between the date of the first referral received and date of last referral removed (or 1 February 2021 if patients had 'open' referrals by 1 February 2021, which was the case in 6 DTD and 512 non-DTD patients), the means and medians were, respectively, 98 and 90 months for the DTD group and 39 and 24 months for the non-DTD group. Again, we excluded 1 DTD and 2 non-DTD patients with a total EHR duration >50 years, as well as 22 DTD and 1033 non-DTD patients with missing or erroneous referral dates.

Socio-demographics

Table 1 provides a summary of socio-demographic characteristics. The DTD group was statistically significantly older than the non-DTD group, although the effect size was very small (52.8 vs 51.6 years, respectively; $p<0.001$; $d=0.06$). Both groups had a majority of females and white patients, with the proportions being larger in the DTD group (61.9% vs 59.1% female ($p<0.001$; $V=0.02$); 82.9% vs 68.8% white ($p<0.001$; $V=0.12$)). In both the DTD and non-DTD groups, those in active employment or studying were a minority, with the proportion being significantly lower in the DTD group (28.2% vs 40.9% ($p<0.001$; $V=0.15$)). The DTD group had a significantly higher proportion of patients in receipt of sickness or disability benefits (28.6% vs 16.2%; $p<0.001$).

Diagnosis and comorbidities

Diagnostic characteristics are summarised in Table 2. While about a third (32.5%) of the patients in the DTD group had been depressed for 5 years or less, another third (33.8%) had been depressed for 9 years or more. In the non-DTD group, 41% of patients had disease duration <5 years and 27% >9 years. Rates of any ICD-10 diagnosis being recorded ranged from approximately 14% to 66% (weighted average 45%) in the NHS Trusts that provided data. There was a higher proportion of recurrent depression (F33.x) in the DTD group compared with the non-DTD group (50.4% vs 30.0% ($p<0.001$; $V=0.17$)). The DTD group also had a higher proportion of patients with episodes of severe depression, with or without psychosis (F32.2, F32.3, F33.2 or F33.3; 34.1% vs 23.6% ($p<0.001$; $V=0.09$)) and episodes of psychotic depression (F32.3 or F33.3; 14.8% vs 10.0% ($p<0.001$; $V=0.06$)) than the non-DTD group. However, the proportion of comorbid acute and transient psychotic disorders (F23.X) was not significantly different between the DTD and non-DTD groups ($p=0.074$) (Table 3). The rate of self-harm and suicidal ideation recorded in the EHRs was significantly higher in the DTD compared with of the non-DTD group, with a moderate-to-large effect size (59.5% vs 13.3% ($p<0.001$; $V=0.43$)).

In both the DTD and non-DTD groups, the most common psychiatric comorbidities were substance misuse, personality disorders, anxiety disorders and adjustment disorders (Table 3). The DTD group had a significantly higher proportion of patients with comorbid substance misuse (18.5% vs 6.9% ($p<0.001$; $V=0.16$)) and personality disorders (14.3% vs 4.9% ($p<0.001$; $V=0.15$)) than the non-DTD group. In both the DTD and non-DTD groups, the most common physical health comorbidities were essential hypertension and type 2 diabetes mellitus, with both being significantly more common in the DTD group, although effect sizes were small ($p<0.001$ and $V<0.1$ for both).

Table 2. Diagnostic characteristics in DTD and non-DTD groups.

	DTD group ($n=5237$)	Non-DTD group ($n=22,947$)	p value[a]	Effect size[b]
Recurrence, *n* (%)			$p<0.001$	
Recurrent depressive disorder (F33.X)	2642 (50.4)	6877 (30.0)		$V=0.17$
Depressive episode (F32.X)	2595 (49.6)	16,070 (70.0)		
Severity, *n* (%)				
Severe depression, with or without psychosis (F32.2, F32.3, F33.2, F33.3)	1788 (34.1)	5425 (23.6)	$p<0.001$	$V=0.09$
Psychotic features, *n* (%)				
Psychotic symptoms, in depressive episode (F32.3) or recurrent depressive disorder (F33.3)	775 (14.8)	2303 (10.0)	$p<0.001$	$V=0.06$
Duration of depression from initial diagnosis, *n* (%)			$p<0.001$	
⩽5 years	1704 (32.5)	9326 (40.6)		
5–7 years	1242 (23.7)	5027 (21.9)		
7–9 years	519 (9.9)	2383 (10.4)		
⩾9 years	1772 (33.8)	6211 (27.1)		
Suicidality, *n* (%)				
Recorded self-harm or suicide attempts	3116 (59.5)	3063 (13.3)	$p<0.001$	$V=0.43$

DTD: difficult-to-treat depression.
[a]Chi-square test with Bonferroni correction for multiple comparisons.
[b]Cramer's V statistic with 1 degree of freedom ($V \approx 0.1$ represents small effect size, $V \approx 0.2$ represents small-to-medium effect size, $V \approx 0.4$ represents medium-to-large effect size).

Table 3. Psychiatric and physical health comorbidities.

ICD-10 codes	DTD group ($n=5237$)	Non-DTD group ($n=22,947$)
Comorbid psychiatric diagnoses, *n* (%)		
F10–F19 – psychoactive substance abuse	971 (18.5)	1589 (6.9)
F10 – alcohol	472 (9.0)	834 (3.5)
F12 – cannabinoids	89 (1.7)	187 (0.8)
F13 – sedatives or hypnotics	38 (0.7)	31 (0.1)
F14 – cocaine	25 (0.5)	54 (0.2)
F17 – tobacco	184 (3.5)	246 (1.1)
F19 – multiple drugs	163 (3.1)	237 (1.0)
F23.X – Acute and transient psychotic disorders	36 (0.7)	110 (0.5)
F34 – Persistent mood disorders	149 (2.8)	209 (0.9)
F40 – Phobic anxiety disorders	151 (2.9)	257 (1.1)
F41 – Other anxiety disorders	636 (12.2)	1293 (5.4)
F42 – Obsessive-compulsive disorder	193 (3.7)	354 (1.5)
F43 – Reaction to severe stress and adjustment disorder	628 (12.0)	1507 (6.3)
F44 – Dissociative disorders	37 (0.7)	56 (0.2)
F45 – Somatoform disorders	44 (0.8)	92 (0.4)
F50 – Eating disorders	193 (3.7)	521 (2.2)
F60 – Specific personality disorders	749 (14.3)	1115 (4.9)
G47 – Sleep disorders	17 (0.3)	34 (0.1)
Comorbid physical health diagnoses, *n* (%)		
E10 – Type 1 diabetes mellitus	22 (0.4)	29 (0.1)
E11 – Type 2 diabetes mellitus	80 (1.5)	105 (0.5)
I10 – Essential hypertension	134 (2.6)	185 (0.8)
I20 – Angina pectoris	12 (0.2)	9 (0.04)
I25 – Chronic ischaemic heart disease	13 (0.2)	25 (0.1)
J45 – Asthma	91 (1.7)	99 (0.4)
M19 – Arthrosis, unspecified	7 (0.1)	22 (0.1)
R52 – Pain, unspecified	23 (0.4)	17 (0.1)

DTD: difficult-to-treat depression; ICD-10: International Classification of Diseases, Tenth Revision.

Table 4. Pharmacological treatments prescribed for mental health conditions.

Patients with prescription, *n* (%)	DTD group (*n* = 5237)	Non-DTD group (*n* = 22,947)	*p* value[a]	Effect size[b]
Drugs for depression – total	5237 (100)	18,259 (79.6)	$p < 0.001$	$V = 0.21$
SSRI	4805 (91.8)	15,450 (67.3)	$p < 0.001$	$V = 0.21$
SNRI	3118 (59.5)	5069 (22.1)	$p < 0.001$	$V = 0.32$
NaSSA	3522 (67.3)	7289 (31.8)	$p < 0.001$	$V = 0.28$
TCA	1687 (32.2)	2717 (11.8)	$p < 0.001$	$V = 0.22$
MAOI	198 (3.8)	207 (0.9)	$p < 0.001$	$V = 0.09$
Other	1053 (20.1)	1300 (5.7)	$p < 0.001$	$V = 0.20$
Combinations of drugs for depression	4596 (87.8)	9289 (40.5)	$p < 0.001$	$V = 0.37$
Augmentation with drugs for psychosis – total	3157 (60.3)	6389 (27.8)	$p < 0.001$	$V = 0.27$
Quetiapine	2036 (38.9)	3219 (14.0)	$p < 0.001$	$V = 0.25$
Olanzapine	1342 (25.6)	2659 (11.6)	$p < 0.001$	$V = 0.16$
Risperidone	872 (16.7)	1560 (6.8)	$p < 0.001$	$V = 0.14$
Aripiprazole	856 (16.3)	1309 (5.7)	$p < 0.001$	$V = 0.16$
Lurasidone	28 (0.5)	26 (0.1)	$p < 0.001$	$V = 0.04$
Other drugs for psychosis	655 (12.5)	882 (3.8)	$p < 0.001$	$V = 0.15$
Augmentations with drugs for relapse prevention ('mood stabilisers')				
Lithium	864 (16.5)	1261 (5.5)	$p < 0.001$	$V = 0.16$
Valproate	135 (2.6)	230 (1.0)	$p < 0.001$	$V = 0.05$
Carbamazepine	154 (2.9)	185 (0.8)	$p < 0.001$	$V = 0.08$
Lamotrigine	219 (4.2)	268 (1.2)	$p < 0.001$	$V = 0.09$
Drugs for insomnia				
Benzodiazepines	3191 (60.9)	6969 (30.4)	$p < 0.001$	$V = 0.25$
Hypnotics (Z-drugs)	2647 (50.5)	5438 (23.7)	$p < 0.001$	$V = 0.23$

DTD: difficult-to-treat depression; MAOI: monoamine oxidase inhibitor; NaSSA: noradrenergic and specific serotonergic antidepressant; SNRI: serotonin and norepinephrine reuptake inhibitor; SSRI: selective serotonin reuptake inhibitor; TCA: tricyclic antidepressant.

[a]Chi-square test with Bonferroni correction for multiple comparisons.

[b]Cramer's *V* statistic with 1 degree of freedom ($V \approx 0.1$ represents small effect size, $V \approx 0.2$ represents small-to-medium effect size, $V \approx 0.3$ represents medium effect size, $V \approx 0.4$ represents medium-to-large effect size).

Treatment

Psychotropic medications identified in EHRs are summarised in Table 4. Compared to the non-DTD group, a higher proportion of patients in the DTD group had had prescriptions for drugs for depression (100% vs 79.6%), augmentation with drugs for psychosis ('antipsychotics', 60.3% vs 27.8%) and lithium augmentation (16.5% vs 5.5%), as well as prescriptions for benzodiazepines (60.9% vs 30.4%) and 'z-drugs' (50.5% vs 23.7%) (all $p < 0.001$).

The majority of patients in the DTD group (83.4%) had been prescribed three or more drugs for depression and 17% had been prescribed two. Consistent with the study DTD group criteria, no patients in the DTD group had been prescribed fewer than two drugs for depression. In the non-DTD group, 50% of the sample had been prescribed two or more drugs for depression. In total, 3780 patients in the non-DTD group (16.5%) had received four or more drugs for depression, but were excluded from the DTD group based on other criteria relating to disease duration/persistence/recurrence.

The most commonly prescribed classes of drugs for depression, both in the DTD and non-DTD groups, were SSRIs (91.8% and 67.3%, respectively), noradrenergic and specific serotonergic antidepressants (NaSSAs) (67.3% and 31.8%, respectively), serotonin and norepinephrine reuptake inhibitors (SNRIs) (59.5% and 22.1%, respectively) and tricyclic antidepressants (TCAs) (32.2% and 11.8%, respectively) (Table 4). Independent of base class of drug for depression, and for both the DTD and non-DTD groups, the most common augmentation strategies were quetiapine (38.9% and 14.0%, respectively), olanzapine (25.6% and 11.6%), risperidone (16.7% and 6.8%), aripiprazole (16.4% and 5.7%) and lithium (16.5% and 5.5%). Benzodiazepines had been prescribed to 60.9% and 30.4% of patients in DTD and non-DTD groups, respectively. Augmentation with thyroid hormones (T3 or T4) was used in few patients in either group (<1% of patients overall). Combination therapy with drugs for depression (summarised in Supplementary Table 2) had been used by 87.8% of patients in the DTD group, compared with 40.5% in the non-DTD group. Only 9.2% of patients in the DTD group had not used combination of drugs for depression or augmentation (only monotherapy with drugs for depression was recorded in their EHRs), compared with 34.3% of the non-DTD group.

The majority of drugs for depression were used at higher doses in the DTD group than the non-DTD group. While dosage differences reached statistical significance for most drugs for depression, effect sizes were generally small (Supplementary Table 3).

Non-pharmacological interventions were used significantly more frequently in the DTD than non-DTD group. In total, 85.8% of patients in the DTD group had psychological therapies recorded in their EHRs, compared with 30.2% in the non-DTD group ($p < 0.001$; $V = 0.44$). The proportions of patients referred for ECT were 16.4% vs 3.8%, respectively ($p < 0.001$; $V = 0.20$).

Table 5. Pharmacological treatments prescribed for physical health conditions.

Patients with prescription, *n* (%)	DTD group (*n* = 5237)	Non-DTD group (*n* = 22,947)	*p* value[a]	Effect size[b]
Statins	971 (18.5)	2033 (8.9)	$p < 0.001$	$V = 0.12$
Antihypertensives	1558 (29.7)	3174 (13.8)	$p < 0.001$	$V = 0.17$
β-blockers	1018 (19.4)	1830 (8.0)	$p < 0.001$	$V = 0.15$
Others	833 (15.9)	1966 (8.6)	$p < 0.001$	$V = 0.16$
Anti-diabetic agents	453 (8.6)	919 (4.0)	$p < 0.001$	$V = 0.08$
Anti-coagulants/Anti-platelets	428 (8.2)	972 (4.2)	$p < 0.001$	$V = 0.07$
Aspirin	265 (5.1)	817 (3.6)	$p < 0.001$	$V = 0.03$
Bronchodilators	589 (11.2)	1012 (4.4)	$p < 0.001$	$V = 0.11$
Analgesics	2443 (46.6)	5406 (23.6)	$p < 0.001$	$V = 0.20$
Opioids	1702 (32.5)	3520 (15.3)	$p < 0.001$	$V = 0.17$
NSAIDs	1555 (29.7)	3175 (13.8)	$p < 0.001$	$V = 0.16$

DTD: difficult-to-treat depression; NSAID, non-steroidal anti-inflammatory drug.
[a]Chi-square test with Bonferroni correction for multiple comparisons.
[b]Cramer's *V* statistic with 1 degree of freedom ($V \approx 0.1$ represents small effect size, $V \approx 0.2$ represents small-to-medium effect size).

Table 5 provides a summary of the pharmacological treatments recorded as being prescribed for physical health conditions at any point in the patients' EHRs. The most common classes of drugs in both the DTD and non-DTD group were analgesics (46.6% vs 23.6%), antihypertensives (29.7% vs 13.8%) and statins (18.5% vs 8.9%). In all cases, the rate of prescription was statistically significantly higher in the DTD group ($p < 0.001$).

Admissions

A total of 728 patients (13.9%) in the DTD group and 1791 patients (7.8%) in the non-DTD group had been admitted to a mental health ward at least once during the period covered by their EHRs. For those patients admitted at least once, the admission rates (per patient per year) were significantly higher in the DTD group (mean 1.88 vs 1.12 admissions ($p < 0.001$; $d = 0.76$)). The admission rate across the whole sample was similarly greater for patients with DTD compared with those with non-DTD (0.27 vs 0.11 admissions per patient per year ($p < 0.001$; $d = 0.23$)). Average length of admission was similar between groups (mean 55.7 vs 61.1 days ($p = 0.61$; $d = 0.015$)), with median values being identical (30 days). In total, 349 patients were excluded from admission rate and length analysis due to implausible recorded admission dates.

Discussion

'Difficult-to-treat depression (DTD)' is a new concept and a proposed definition is 'depression that continues to cause significant burden despite usual treatment efforts' (McAllister-Williams et al., 2020). This is intentionally not a rigidly objective definition, but rather one to be used clinically, reflecting the situation jointly faced by patient and clinician. When a patient's depression is described as DTD will vary between clinicians and health care settings. A consequence is that prevalence will vary according to setting (both geographically and level of care). However, this is the first study that has attempted to obtain an estimate of the prevalence of DTD, using proxy measures that suggest patient and clinician may perceive the clinical situation as being challenging. These proxy measures are a mixture of evidence of a prolonged episode of depression and/or multiple episodes, combined with a history of use of multiple treatments with drugs for depression. These criteria used for possible/probable DTD have been applied to electronic clinical records from five large secondary-care mental health Trusts in the United Kingdom, including the use of an NLP algorithm to identify data from free text entries, a rule-based algorithm for data related to psychotherapy, ECT and suicidality and a deep learning NLP model to extract unstructured data on pharmacological treatments. Of those adults identified with depression, 19% met the criteria for possible/probable DTD. They were characterised by being more likely to have episodes of severe depression (with and without psychosis), have recorded suicide attempts or incidents of self-harm, suffer from psychiatric and/or physical health comorbidities, be admitted to a mental health ward and be unemployed or on sickness or disability benefits. They also had higher use of all modalities of treatment.

There is significant variation in the literature regarding the definition of TRD (Brown et al., 2019). A Delphi-method-based consensus on the definition of TRD for clinical trials (Sforzini et al., 2021) strongly supported the notion of two failed treatments as a relevant threshold. There was moderate consensus that these two treatments might include neurostimulation treatments such as ECT. The same guidelines recommend that failed courses of psychotherapy should not be considered as one of the previous treatments required to meet criteria for TRD, but with some disagreement between experts (Sforzini et al., 2021). Our sample was identified in secondary, specialist, mental health care. Data from UK primary care suggest that accessing specialist mental health services can be challenging (Telford et al., 2002) and only a small proportion of patients presenting with TRD in primary care are referred to secondary care (Wiles et al., 2018). Therefore, it could be argued that our DTD sample captured a particularly unwell/difficult-to-treat group of patients compared with many conventionally defined 'TRD' populations. The DTD framework is based on expert consensus, is patient-centred and the definition of what constitutes 'significant burden' is subjective (McAllister-Williams et al., 2020, 2021a). The commonly used TRD threshold of two previous treatment courses with drugs for depression

was not sufficient to define our DTD group; we required at least four previous antidepressive treatments, two of which could be psychological therapies or ECT, meaning that all our DTD samples met a TRD definition based on two failed drugs for depression. This arguably makes our definition of DTD both more inclusive – by including non-pharmacological treatments – but also more demanding, by essentially doubling the number of treatment failures required in most TRD definitions. We also included markers of chronicity, such as the current depressive episode lasting ⩾2 years or ⩾3 depression episodes in the totally of the EHR. This recognises the difficulty in distinguishing discrete episodes and reflects the chronic, waxing and waning nature of depression for many patients.

One key measure of burden of illness is the amount of time individuals are engaged in secondary-care specialist mental health services. Unsurprisingly, this was higher in the DTD group. Assessing duration of active contact from EHRs is complex due to the imprecise nature of defining and identifying when a person is under 'active' care. When utilising the information in the EHRs regarding dates of referrals into, and discharge from, various elements of services, patients defined as having DTD had much longer durations of contact with services (median 90 months) compared with non-DTD patients (median 24 months). This is despite the observation that the duration of EHR from the first entry in 'progress notes' to our census date of 1 February 2021 (or death) was not so different between the two groups (medians of 121 months for DTD group and 104 months for non-DTD group). This is in agreement with the DTD group having more active involvement over a not dissimilar period of time from first contact, compared with the non-DTD group, consistent with a higher burden of illness.

The ICD-10 diagnostic codes showed that the DTD group had more cases of severe and recurrent depression (although the effect sizes were small). This is not surprising given the selection criteria. The data collected from EHRs did not include other measures of disease severity post-diagnosis (e.g. mood rating scales) or disease course (such as frequency of relapse). We do have indicators of risk and admission frequency, which could be construed as proxies for severity. More than 10% of the patients in the non-DTD group had a history of suicidal attempts or self-harm. This suggests high levels of distress and risk, albeit much lower than the almost 60% of patients in the DTD group with a history of suicidality or self-harm. As this is a cross-sectional study, we cannot comment on causality, but it is possible that risk to self was a driver for referral to secondary care, more aggressive treatment and/or admission to mental health wards. Indeed, the admission rates (per patient per year) were significantly higher in the DTD group.

The depressed patients making up the non-DTD group (81% of our sample) should not, however, be assumed to be straightforward ('easy-to-treat') or 'low-risk' cases. Approximately three-quarters of this group had had an episode of depression lasting ⩾2 years or had ⩾3 episodes of illness in the totally of the EHR. We should not ignore the fact that these patients were unwell enough (or there were sufficient concerns regarding safety) to be referred to secondary care. They may have failed to meet the DTD criterion of ⩾4 antidepressive treatments not because they did not need more intensive treatment, but simply because it had not been offered (thus far) to non-DTD patients, who had generally been under secondary care for a relatively short period (median 11 months, compared with 69 months in the DTD group). Potential barriers to access to secondary care and adequate treatment warrant further investigation. It should be noted that 50% of the non-DTD group had been prescribed two or more drugs for depression and would therefore be commonly classified at TRD. It seems likely that with additional follow-up time a relevant proportion of patients in the non-DTD group would be classed as DTD. The non-DTD group included a higher proportion of non-white patients and males than the DTD group; if engagement with treatment is lower for some demographic groups than others, this is an important issue to address to ensure equity of patient access to relevant treatment pathways.

Rates of comorbid illness, both mental and physical, were higher in the DTD than the non-DTD group. The presence of significant comorbidity in the DTD group may constitute one aspect of the difficulty of treating these patients; for example, treatment options may be more limited for patients with comorbid heart disease or diabetes, who may not be able to tolerate medication classes such as monoamine oxidase inhibitors (MAOIs), TCAs or drugs for psychosis with more marked metabolic or cardiac side effects. The relationship between DTD and comorbidities warrants further investigation, to determine whether DTD increases the risk of some comorbid conditions, whether some comorbid conditions may be underlying causal factors in the development of DTD or whether there may be shared causal mechanisms. Regardless of a causal relationship being present or not, identifying and treating comorbidities is an important element of the management of DTD (McAllister-Williams et al., 2020).

A core element of defining DTD is that 'usual treatment efforts' have been undertaken (McAllister-Williams et al., 2020) – we sought to determine what treatment efforts had been attempted in 'real-world' secondary care in the United Kingdom and could therefore be labelled 'usual'. Our analysis revealed greater intensification of treatment strategies in the DTD group compared with the non-DTD group. There was a consistent trend for drugs for depression to be used at higher dosages in the DTD group, although differences in mean dosages were generally small. Most patients in the DTD group (88%) had received combination therapy with drugs for depression at some point, while only 9% had received neither combination nor augmentation treatment. Combinations of drugs for depression were used at a much lower rate in the non-DTD group (41%), while one-third of non-DTD patients (34%) had received only monotherapy. This surprisingly high rate of combination therapy may reflect a limitation of the Med7 NLP model: although it was designed to determine the nature of 'current' drug 'mentions' as far as possible, scope for misinterpretation remains. For example, two drugs for depression can be mentioned in one entry to record the rationale for selection of one drug over another; the NLP model might pick up both drug names and default to an assumption that both were prescribed. However, this potential over-representation of prescribed drugs would be expected to affect DTD and non-DTD groups equally; the relative difference between groups therefore likely reflects a genuine difference in the rate of use of combination therapy, even if absolute prescription rates might be lower for both groups.

In terms of specific treatments, SSRIs were the most commonly used class of drugs for depression in both DTD and non-DTD groups, followed by NaSSAs and SNRIs. This is broadly in

line with UK guidelines for antidepressant pharmacological therapy (Cleare et al., 2015; NICE, 2009). The relatively high rate of NaSSA prescriptions reflects use of mirtazapine in combination with other drugs for depression (Supplementary Table 2); SNRIs were also used largely as part of combination therapy.

Treatment differences between the groups were to be expected, since number of antidepressive treatments formed part of the criteria for defining the DTD group, who also had longer durations of contact with services. Nonetheless, our findings demonstrate use of more intensive strategies, including combination therapy, augmentation and non-pharmacological interventions, in patients categorised as having DTD. Encouragingly, this is in line with treatment guidelines that recommend a range of treatment options beyond monotherapy with drugs for depression (Cleare et al., 2015; NICE, 2009).

Augmentation, most commonly with drugs for psychosis but also with drugs for relapse prevention ('mood stabilisers'), was much more frequently used in the DTD group than the non-DTD group, as were non-pharmacological interventions. Quetiapine was the most common augmentation choice, consistent with guidelines (Cleare et al., 2015; NICE, 2009). Olanzapine was the next most common augmentation agent, despite only being a second-line recommendation in most guidelines (Cleare et al., 2015; Taylor et al., 2020). Aripiprazole was used at a lower rate and may be underused considering it is an effective augmentation strategy with an excellent metabolic side-effect profile and dopaminergic effects (Berman et al., 2009). We do not have information on the temporal sequence in which any of the drugs for depression or augmentation treatment options were trialled, and first/early treatment choices prescribed under primary care would not be captured in the EHRs.

Non-pharmacological interventions showed a marked difference between DTD and non-DTD groups in our study. This may in part be an artefact of the selection criterion based on number of treatments, as non-pharmacological treatments could contribute to the four or more treatments that determined DTD status. Only 30% of the non-DTD group had had psychotherapy, compared with 86% of the DTD group, suggesting that psychotherapy may be used at a later stage of the treatment pathway in many patients. It is also possible that psychotherapy provided via primary care and completed prior to referral to secondary care (such as Improving Access to Psychological Therapies (IAPT)) was not captured in our analysis. Perhaps less surprising was that the group with fewer (<4) treatments overall had rarely received ECT (4% of the non-DTD group), as multiple treatment failures are likely to be prerequisite to progressing to what could be perceived as more invasive treatment. However, use of ECT was still low (16%) in DTD group. We do not have data on numbers of patients who were referred to tertiary care for more specialist treatment; for example, alternative neurostimulatory options such as vagus nerve stimulation (VNS) that may have a place in some patients with DTD (McAllister-Williams et al., 2021b).

There are a number of limitations to this data set and our analysis. Our identification of patients was reliant on ICD-10 diagnoses being recorded and correct, and on average these were only recorded in 45% of the overall patient populations of the NHS Trusts providing data. Our sample therefore reflects a small proportion of patients in secondary psychiatric care in these Trusts, and may be skewed towards patients with admissions, since diagnosis fields may be more likely to be completed on admission or discharge from a ward. In addition, if an excluding diagnosis was not recorded, then patients could have been erroneously included. Trusts in Southern England and London are overrepresented in our sample, which is therefore not representative of the United Kingdom overall.

All EHRs as of 1 February 2021 of patients meeting the inclusion criteria were included, irrespective of whether the patient was alive or not. Determining which of the patients included were actively engaged in services as of 1 February 2021 was not possible. Patients can technically be in services for long periods with no entries made in their notes. Referral and discharge details are provided for individual service elements rather than specialist care in its entirety and some patients have not apparently been discharged and yet have had no new entries made in their EHR for years. These issues present a challenge with regard to calculating the duration of a patient's EHR. We utilised two methods, one relying on accurate recording of referral into and discharge from elements of specialist services, and the other entire duration of the EHRs, bearing in mind that EHRs are never closed. The first method provides some degree of estimate of duration of contact with services, while the second estimates the period of time since a patient had first contact.

Lack of longitudinal and post-diagnosis illness severity data, and the nature in which referrals and discharges were recorded, makes identification of number and duration of episodes challenging. We have tried to address this using a robust operational definition. However, given the limitations of our approach, some patients may have incorrectly been classified as suffering from non-DTD rather than DTD, and to a probable lesser extent, vice versa. Furthermore, the treatment history dimension of our DTD criteria may have led to a non-DTD classification for some patients who were clinically similar to those in the DTD group, but for some reason had been offered fewer treatments. Our estimate of 19% of MDD patients therefore likely underestimates the true prevalence of DTD within the secondary care setting. It is also worth acknowledging the inherent difficulty in attempting to dichotomise patients within what is in reality a complex, multidimensional continuum (McAllister-Williams et al., 2020). While we have labelled the patient groups as 'DTD' and 'non-DTD' for simplicity, we recognise that many patients falling into the 'non-DTD' category in fact have a degree of difficulty associated with their treatment.

The study was a retrospective analysis based on unstructured data mining of a convenience sample. NLP modelling provides an opportunity to harness extensive data sets from unstructured EHRs, but there are limitations to the accuracy with which data are extracted, for example, the potential issue of overcounting medication prescriptions as described above. Nonetheless, previous studies using the same UK-CRIS system have demonstrated the feasibility of this approach (Goodday et al., 2020; Kormilitzin et al., 2021; Vaci et al., 2020). Also, the inclusion of a review of clinical documentation in the assessment of the efficacy of antidepressant treatment trials has been highlighted as a recommendation for practice (Sforzini et al., 2021). In interpreting the data extracted in an analysis of this type, one should be mindful that patterns and relative proportions are more meaningful than the precise absolute numerical values. A further potential source of error is that some aspects of the raw data to which the NLP algorithm is applied are likely to be less than optimal. For example, rates of physical health screening and diagnosis of

physical illnesses is known to be sub-optimal in patients with mental illness (Pearsall et al., 2019), and a high rate of missing data has been observed for social and behavioural parameters in EHRs (Goodday et al., 2020). There is also the potential of 'siloing', with physical health services operating in different organisations from those where mental health care is being provided. These factors are likely to have led to a large under-representation of the rates of physical comorbidities in our sample. Finally, our analysis showed increased health care resource utilisation among patients with DTD compared with 'non-DTD' MDD; estimating the economic burden associated with this, as well as the socio-economic impacts of DTD, was beyond the scope of this analysis, but is an important area for future research.

Conclusion

This study demonstrates that it is possible to identify individuals suffering from possible/probable DTD from electronic case records using unstructured data mining. Such patients have a strikingly higher burden of illness, as indexed by number and severity of episodes, suicidality, rates of comorbidity and employment status, as well as rates of hospital admission and duration of active involvement with specialist services. This burden of illness is likely to be associated with higher direct and indirect economic costs. As such, this group of patients warrant both identification but also utilisation of the model of care for DTD proposed by an international consensus group, including reviewing of diagnosis, identification and treatment of comorbidities, identification of any aetiological factors that may be tractable and then use of all modalities of treatment available. Given the high prevalence of depression and its burden, investment in the DTD pathway is likely to lead to benefit across society.

Acknowledgements

The authors thank David Newton (Akrivia Health, Oxford) for his contribution to the analyses, and Emily Haworth (Newcastle University) and Samantha Stanbury, PhD (an independent medical writer contracted to Vivari Communications Ltd, and funded by LivaNova PLC) for contributions to manuscript drafts. We also thank Simon Walker (Cumbria, Northumberland, Tyne and Wear NHS Foundation Trust), Samantha Scholtz (West London NHS Trust), Heloise Mongue-Din (South West London & St George's Mental Health Trust), Tanya Smith and Suzanne Fisher (Oxford Health NHS Foundation Trust) and Peter Phiri (Southern Health NHS Foundation Trust) for their assistance with obtaining data from electronic health records.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship and/or publication of this article: Tiago Costa declares no financial conflicts of interest. Bayar Menzat, Tomas Engelthaler, Benjamin Fell, Tarso Franarin, Gloria Roque, Yiran Wei and Xinyue Zhang are current or former employees of Akrivia Health, which received funding from LivaNova PLC for the conduct of this analysis. Hamish McAllister-Williams has received fees from American Center for Psychiatry & Neurology United Arab Emirates, British Association for Psychopharmacology, European College of Neuropsychopharmacology, International Society for Affective Disorders, Janssen, LivaNova, Lundbeck, My Tomorrows, OCM Comunicaziona s.n.c., Pfizer, Qatar International Mental Health Conference, Sunovion, Syntropharma, UK Medical Research Council and Wiley; grant support from National Institute for Health Research Efficacy and Mechanism Evaluation Panel and Health Technology Assessment Panel; and non-financial support from COMPASS Pathways and MagStim.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The study and analysis were commissioned and funded by LivaNova PLC.

ORCID iD

Tiago Costa https://orcid.org/0000-0002-3666-7531

Supplemental material

Supplemental material for this article is available online.

References

Berman RM, Fava M, Thase ME, et al. (2009) Aripiprazole augmentation in major depressive disorder: A double-blind, placebo-controlled study in patients with inadequate response to antidepressants. *CNS Spectrums* 14(4): 197–206.

Brown S, Rittenbach K, Cheung S, et al. (2019) Current and common definitions of treatment-resistant depression: Findings from a systematic review and qualitative interviews. *Canadian Journal of Psychiatry* 64(6): 380–387.

Cai W, Mueller C, Li YJ, et al. (2019) Post stroke depression and risk of stroke recurrence and mortality: A systematic review and meta-analysis. *Ageing Research Reviews* 50: 102–109.

Cleare A, Pariante CM, Young AH, et al. (2015) Evidence-based guidelines for treating depressive disorders with antidepressants: A revision of the 2008 British Association for Psychopharmacology guidelines. *Journal of Psychopharmacology* 29(5): 459–525.

Dome P, Kunovszki P, Takacs P, et al. (2021) Clinical characteristics of treatment-resistant depression in adults in Hungary: Real-world evidence from a 7-year-long retrospective data analysis. *PLoS ONE* 16(1): e0245510.

Feng L, Li L, Liu W, et al. (2019) Prevalence of depression in myocardial infarction: A PRISMA-compliant meta-analysis. *Medicine* 98(8): e14596.

Gagnon LM and Patten SB (2002) Major depression and its association with long-term medical conditions. *Canadian Journal of Psychiatry* 47(2): 149–152.

Gaynes BN, Lux L, Gartlehner G, et al. (2020) Defining treatment-resistant depression. *Depress and Anxiety* 37(2): 134–145.

Goodday SM, Kormilitzin A, Vaci N, et al. (2020) Maximizing the use of social and behavioural information from secondary care mental health electronic health records. *Journal of Biomedical Informatics* 107: 103429.

Greenberg PE, Fournier AA, Sisitsky T, et al. (2015) The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *Journal of Clinical Psychiatry* 76(2): 155–162.

Johnson AE, Pollard TJ, Shen L, et al. (2016) MIMIC-III, a freely accessible critical care database. *Scientific Data* 3: 160035.

Johnston KM, Powell LC, Anderson IM, et al. (2019) The burden of treatment-resistant depression: A systematic review of the economic and quality of life literature. *Journal of Affective Disorders* 242: 195–210.

Kormilitzin A, Vaci N, Liu Q, et al. (2021) Med7: A transferable clinical natural language processing model for electronic health records. *Artificial Intelligence in Medicine* 118: 102086.

Li G, Fife D, Wang G, et al. (2019) All-cause mortality in patients with treatment-resistant depression: A cohort study in the US population. *Annals of General Psychiatry* 18: 23.

McAllister-Williams RH, Arango C, Blier P, et al. (2020) The identification, assessment and management of difficult-to-treat depression:

An international consensus statement. *Journal of Affective Disorders* 267: 264–282.

McAllister-Williams RH, Arango C, Blier P, et al. (2021a) Reconceptualising treatment-resistant depression as difficult-to-treat depression. *The Lancet Psychiatry* 8(1): 14–15.

McAllister-Williams RH, Bulmer S, Newton K, et al. (2021b) Assessment for vagus nerve stimulation in patients with difficult-to-treat depression: A model from the Newcastle Regional Affective Disorders Service (RADS). *Journal of Affective Disorders* 280(Pt. A): 315–318.

McAllister-Williams RH, Christmas DMB, Cleare AJ, et al. (2018) Multiple-therapy-resistant major depressive disorder: A clinically important concept. *British Journal of Psychiatry* 212(5): 274–278.

McCrone P, Dhanasiri S, Patel A, et al. (2008) *Paying the Price: The Cost of Mental Health Care in England to 2026*. London: King's Fund.

Malhi GS and Mann JJ (2018) Depression. *The Lancet* 392(10161): 2299–2312.

Martin-Cook K, Palmer L, Thornton L, et al. (2021) Setting measurement-based care in motion: Practical lessons in the implementation and integration of measurement-based care in psychiatry clinical practice. *Neuropsychiatric Disease and Treatment* 17: 1621–1631.

NICE (2009) *Depression in Adults: Treatment and Management – Clinical Guideline [CG90]*. London: NICE.

Nouwen A, Winkley K, Twisk J, et al. (2010) Type 2 diabetes mellitus as a risk factor for the onset of depression: A systematic review and meta-analysis. *Diabetologia* 53(12): 2480–2486.

Pearsall R, Shaw RJ, McLean G, et al. (2019) Health screening, cardiometabolic disease and adverse health outcomes in individuals with severe mental illness. *BJPsych Open* 5(6): e97.

Read JR, Sharpe L, Modini M, et al. (2017) Multimorbidity and depression: A systematic review and meta-analysis. *Journal of Affective Disorders* 221: 36–46.

Reutfors J, Andersson TM, Brenner P, et al. (2018) Mortality in treatment-resistant unipolar depression: A register-based cohort study in Sweden. *Journal of Affective Disorders* 238: 674–679.

Rush AJ, Aaronson ST and Demyttenaere K (2019) Difficult-to-treat depression: A clinical and research roadmap for when remission is elusive. *Australian and New Zealand Journal of Psychiatry* 53(2): 109–118.

Rush AJ, Sackeim H, Conway C, et al. (2022) Clinical research challenges posed by difficult-to-treat depression. *Psychological Medicine* 52: 419–432.

Rush AJ, Trivedi MH, Wisniewski SR, et al. (2006) Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *American Journal of Psychiatry* 163(11): 1905–1917.

Rush AJ, Zimmerman M, Wisniewski SR, et al. (2005) Comorbid psychiatric disorders in depressed outpatients: Demographic and clinical features. *Journal of Affective Disorders* 87(1): 43–55.

Russell JM, Hawkins K, Ozminkowski RJ, et al. (2004) The cost consequences of treatment-resistant depression. *Journal of Clinical Psychiatry* 65(3): 341–347.

Sforzini L, Worrell C, Kose M, et al. (2021) A Delphi-method-based consensus guideline for definition of treatment-resistant depression for clinical trials. *Molecular Psychiatry*. Epub ahead of print 16 December. DOI: 10.1038/s41380-021-01381-x.

Taylor RW, Marwood L, Oprea E, et al. (2020) Pharmacological augmentation in unipolar depression: A guide to the guidelines. *International Journal of Neuropsychopharmacology* 23(9): 587–625.

Telford R, Hutchinson A, Jones R, et al. (2002) Obstacles to effective treatment of depression: A general practice perspective. *Family Practice* 19(1): 45–52.

Trivedi MH, Rush AJ, Wisniewski SR, et al. (2006) Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *American Journal of Psychiatry* 163(1): 28–40.

Vaci N, Liu Q, Kormilitzin A, et al. (2020) Natural language processing for structuring clinical text data on depression using UK-CRIS. *Evidence-Based Mental Health* 23(1): 21–26.

WHO (2017) *Depression and Other Common Mental Disorders: Global Health Estimates*. Geneva: WHO.

Wiles N, Taylor A, Turner N, et al. (2018) Management of treatment-resistant depression in primary care: A mixed-methods study. *British Journal of General Practice* 68(675): e673–e681.