**Abstract**

**Background:** An earlier evaluation (Fox et al., 2014) highlighted reductions in risk behaviours and restrictive practices for women admitted to low secure Dialectical Behaviour Therapy (DBT) unit. Since then, a value-based healthcare model has been adopted. **Aims:** To explore changes in health, social and psychological functioning, risk, quality of life, and in incidents of violence and restrictive practices, over the initial 12-month period of admission to a specialist DBT service. **Method:** Data was extracted from electronic clinical records for 41 women with Emotionally Unstable Personality Disorder admitted to a specialist integrated practice unit (IPU) providing a comprehensive DBT programme. Secondary analysis was conducted on an anonymous dataset of routinely collected outcome measures at baseline admission, and 6 months and 12 months post-admission. ANOVAs and pairwise post hoc comparisons, and non-parametric equivalents, were conducted to examine changes in outcomes. **Results:** Findings showed statistically significant improvements in mental health scores on the ReQOL (p<.01), global, wellbeing, problems, functioning and risk scores on the CORE-OM (allp<.01), and severe disturbance, emotional wellbeing, socioeconomic status, risk and need scores on the HoNOS-Secure (all p<.05). Significant reductions in risk behaviours (p<.01) and restrictive practices (p<.01) were also apparent. The most substantiative improvements were largely demonstrated over a 12-month admission period. **Conclusions:** Admission to the DBT IPU yielded significant improvements on outcomes pertaining to quality of life, psychological distress, and risk. Importantly, these are outcomes that aligned with patients’ perceptions of recovery.

*Keywords:* dialectical behaviour therapy (DBT), Emotionally Unstable Personality Disorder (EUPD), integrated practice unit (IPU), value-based healthcare, outcomes

**An Evaluation of Value-Based Outcomes for Women Admitted to a Dialectical Behaviour Therapy Integrated Practice Unit: A Follow-Up Study**

Emotionally Unstable Personality Disorder (EUPD) is an enduring psychiatric disorder characterized by emotional dysregulation, cognitive distortions, impulsivity and relationship instability (World Health Organisation [WHO], 2019). A variety of psychotherapeutic approaches have been utilised for the treatment of such a complex symptom profile (Storebø et al., 2020), yet the approach currently listed as the recommended treatment for EUPD within clinical guidelines (National Institute for Health and Care Excellence, 2009), is Dialectical Behaviour Therapy (DBT; Linehan et al., 1991). DBT equips patients with skills in managing intense emotions and changing maladaptive coping behaviours through four core strategies of mindfulness, distress tolerance, interpersonal effectiveness, and emotion regulation. A comprehensive DBT programme meets five key functions: enhancing capabilities in the four core strategies through skills training groups; motivating patients within individual therapy; structuring the patient environment; generalizing skills to natural environments through coaching; and enhancing therapists’ capabilities and motivation through team consultations (Linehan, 1993).

The efficacy of DBT for EUPD has been the subject of considerable investigation across a range of inpatient and outpatient settings (see Fox, 2018), though less attention has been paid to secure psychiatric settings. Given that users of such services often present with the most complex, comorbid treatment needs, establishing the effectiveness of treatment approaches is of particular priority in this population. Of the limited research exploring outcomes yielded by DBT for secure psychiatric populations, there is significant variation in the format and structure of DBT programmes being evaluated (see Tomlinson, 2018), which creates difficulty in generalising findings across such settings. Nevertheless, reductions in various risk-related behaviours and clinical outcomes have been documented in individuals with EUPD (Evershed et al., 2003; Moulden et al., 2020).

Despite these initial findings, the effectiveness of DBT in delivering improvements in meaningful, patient-valued outcomes remains a key question. For example, an earlier study by Fox et al. (2014), which evaluated a comprehensive DBT programme for women with EUPD admitted to a specialist low-secure unit, highlighted significant improvements in EUPD symptomology, social functioning, clinical problems and risk behaviours, primarily within the first six months of treatment. However, the outcome measures had not been selected for the DBT service specifically, and thus may not have been important recovery indicators for those accessing the DBT programme. As a result, the effectiveness of the service in delivering improvement in meaningful, patient-valued outcomes remains a key question.

Outcomes have long been posited as a fundamental indicator of the effectiveness and quality of a healthcare system (Donabedian, 1966), though reflection of this at a systems-level is a much more recent development. Traditionally, healthcare systems have operated under *fee-for-service* models in which payment is based upon headcount and number of services provided (Zuvekas & Cohen, 2016). Nevertheless, the move towards an evidence-based and person-centred approach to service provision, as outlined in the ‘Five Year Forward for Mental Health’ guidance (NHS England and NHS Improvement, 2016), has demanded a shift in the way that healthcare services operate. An ‘outcome-based’ agenda reflects such an approach to the organisation and delivery of patient-centred care, recognising clinical outcomes as an imperative in incentivizing the delivery of high quality healthcare (Wallang et al., 2018).

The importance of systematic patient-centred outcome measurement as an indicator of quality of care is fundamental to a true value-based healthcare system (Porter & Lee, 2013). A value-based model is principally founded in the idea that ‘value’ for patients, defined as the sum of health outcomes by the cost of their delivery (Porter, 2010), is the central tenet to an effective healthcare system. However, the outcomes prioritized within this model are ‘patient-derived’, and the focus is on markers which those using the service perceive to be important indicators for their recovery. Given that no standardised benchmark exists for which recovery can be measured against, it is crucial that progress be determined within an individualised, person-centred context.

Since Fox et al.’s (2014) evaluation of a comprehensive women’s DBT programme, the service has undergone a transition, moving towards a value-based healthcare model. As part of this evolution, the outcomes utilised by services were modified and selected through a process of patient consultation (Wallang et al., 2018). The current study was conducted as a follow-up to the previous evaluation of the DBT service, exploring changes in outcomes which those accessing the service consider to be valuable indicators of recovery. Additionally, whereas the initial evaluation explored outcomes for patients admitted to a singular specialist rehabilitation ward, the current study assessed outcomes for patients admitted to the DBT IPU, an integrated practice unit providing a progressive care pathway, with one low-secure unit and two specialist rehabilitation units.

In summary, the present study aimed to explore changes in a series of patient-centred, value-based outcomes over a 12-month admission period for women admitted to the DBT IPU. Specifically, we assessed for changes on patient- and clinician-rated outcome measures of health and social functioning, risk, quality of life and psychological functioning, as well as in incidents of risk behaviours and the use of restrictive practices.

**Method**

**Design**

The current study employed a retrospective within-groups design to evaluate treatment outcomes in women admitted to the DBT IPU between July 2017 and December 2019, over a 12-month admission period. Data was extracted from electronic clinical records and used to build an anonymous database comprising patients’ demographic information, primary ICD-10 (WHO, 1992) diagnosis, level of security and legal status, as well as scores on a range of outcome measures, and incidents of risk behaviours and restrictive practices.

**Participants and Setting**

Overall, 101 women were admitted to the DBT IPU between July 2017 and December 2019. The service is comprised of one low-secure unit and two specialist rehabilitation units and provides inpatient treatment for women detained under the Mental Health Act (MHA; Department of Health, 1983 amended 2007) with a diagnosis of EUPD and additional complex mental health needs through a comprehensive DBT programme, as previously described by Fox et al. (2014). Patients were selected for inclusion if they had completed at least one outcome assessment at baseline, 6-months, and 12-months post-admission. This resulted in a final sample of 41 patients with a complete dataset for at least one outcome.

**Measures**

**Recovering Quality of Life (ReQoL-20).**

The ReQoL-20 (Keetharuth et al., 2018) is a 21-item patient-reported outcome measure designed for use across mental health populations. Items relate to the themes of ‘belonging and relationships’, ‘hope’, ‘activity’, ‘self-perception’, ‘choice, control and autonomy’ and ‘well-being’, with an additional item relating to physical health. Patients are instructed to rate each item in relation to their thoughts, feelings and activities during the last week on a 5-point Likert scale where 0 indicates ‘none of the time’ and 4 indicates ‘most or all of the time’. Possible scores on the mental health domain of this measure range from 0 and 80, with a higher score indicating better quality of life. The physical health domain comprises one item only, and thus possible scores range from 0 to 4, where 0 indicates ‘no problems’ and 4 indicates ‘very severe problems’. The ReQoL-20 has established internal consistency, reliability, validity and sensitivity to change in inpatient mental health populations, including those with EUPD (Keetharuth et al., 2017, 2020).

**Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE-OM).**

The CORE-OM(Evans et al., 2000) is a 34-item patient-reported outcome measure exploring four distress-related domains; wellbeing (4 items), problems (12 items), functioning (12 items) and risk (6 items). Items are rated, based upon distress over the last week, on a 5-point Likert scale where 0 indicates ‘not at all’ and 4 indicates ‘most of the time’, with a higher rating indicating greater problems or distress. Scores for all four domains can be summed to comprise a measure of ‘global distress’; possible scores range from 0 to 136 with a higher score indicating a greater level of global distress. The CORE-OM has established reliability, validity and sensitivity to change and is acceptable for use in a diverse range of clinical and forensic settings (Evans et al., 2002; McCloskey, 2001).

**Health of the Nation Outcomes Scales for Users of Secure and Forensic Services (HoNOS-Secure).**

The HoNOS (Wing et al., 1998) is a 12-item clinician-rated tool for the assessment of health and social functioning. The HoNOS-Secure (Sugarman & Walker, 2007) is a modified version for use in secure and forensic settings. This tool comprises the original twelve HoNOS items assessing behaviour, impairment, symptoms and social functioning, as well as an additional seven ‘security scales’ which assess an individual’s need for a secure environment. Items are scored on a 5-point Likert scale, where 0 indicates ‘no problem’ and 4 indicates ‘severe to very severe problems’. The original twelve items are scored in relation to the most severe presenting problems occuring within the past two weeks. The security items are rated based upon current need for secure care, in consideration of previous behaviours and attitudes, current progress, and potential problems in the coming weeks and months.

For the purpose of this evaluation, scores were calculated in line with an alternative structure outlined by Maddison et al. (2016). This four-factor model divides the original twelve scales into the four subscales of severe disturbance, personal wellbeing, emotional wellbeing and socio-economic status (see Table 1). This alternative structure has been shown to provide a better fit when working with inpatient populations and it has been successfully used to assess progress in secure mental health (Girardi et al., 2018). The seven security scales were also grouped into two subscales based upon the two-factor structure outlined by Tiffin et al. (2011); those assessing ‘risk’ and those assessing ‘need’ (see Table 1). Internal consistency, reliability, validity and sensitivity to change has been established for the HoNOS-secure (Dickens et al., 2007; Pirkis et al., 2005).

**Table 1.** *Modified HoNOS subscales and their associated items*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Insert Table 1 here

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Risk behaviours and restrictive practices.**

Data pertaining to the frequency and severity of risk behaviours (self-harm, violence against persons and objects) and the frequency of the use of restrictive practices (restraints, seclusions)during admissionwas routinely recorded on patients’ electronic records by clinical staff. For the purpose of exploring changes in overall violence, accounting for both the frequency and severity of incidents, aggregated aggression scores (AAS; Alderman et al., 2011; Kay et al., 1988) were also calculated. The ‘weighted’ severity for each incident was first calculated based upon the type of incident (self-harm, violence against persons, violence against objects) and the severity (1 = ‘no harm’, 2 = ‘low’, 3 = ‘moderate’, 4 = ‘serious incident’). To account for aggression type, the severity ratings for incidents of physical aggression against objects, self-harm, and physical aggression against others were multiplied by two, three and four, respectively. To calculate the AAS for each patient, the mean weighted severity of incidents was multiplied by the total number of incidents at each timepoint, separately.

**Data Analysis**

Data was entered into IBM SPSS Statistics version 26. Where the assumption of normality was met, a series of repeated measures ANOVAs and pairwise post hoc comparisons were conducted to test for changes in outcomes across all timepoints (baseline vs. 6 months; 6 months vs. 12 months; baseline vs. 12 months). Where the assumption of normality was not met, the non-parametric Friedman’s test and Wilcoxon Signed Ranks test were conducted. The Bonferroni correction was applied to adjust for multiple comparisons (p < .016). As data was not available on each outcome for the whole sample, the number of patients included in the analyses varied between each outcome.

For statistically significant results, the appropriate effect size was calculated. For repeated measures ANOVAs, partial eta squared ($η^{2}$**p)** was calculated and interpreted based upon the guidelines developed by Cohen (1988): small effect size = 0.01; medium effect size = 0.06; large effect size = 0.14. For paired comparisons, *r* was calculated and interpreted in line with Cohen’s (1988, 1992) guidelines: 0.1 = small effect size, 0.3 = medium effect size, 0.5 = large effect size. Where a non-parametric Friedman’s test was used, *r* was calculated and interpreted in line with Cohen’s (1988, 1992) guidelines for post-hoc comparisons only (Field, 2013).

**Ethical Considerations**

The current study was an evaluation of the DBT service, based on routinely collected clinical assessments, and thus formal ethical approval was not required. However, approval was sought from the organisations internal Clinical Audit and Assurance committee and the services’ clinical director.

**Results**

**Sample Characteristics**

The majority of patients were admitted for treatment to a specialist rehabilitation unit, under Section 3 of the MHA (Department of Health, 1983 amended 2007). All participants had a primary (n=40, 97.6%) or secondary (n=1, 2.4%) diagnosis of EUPD. The age of patients ranged from 18 to 51 years (M=27.56 years, *SD=9.25*). Of the participants included in the sample, 18 had been discharged from the service at the point of data extraction. For these participants, length of stay ranged from 12 to 28 months, with a mean admission length equivalent to 18 months (M=551.36, *SD=148.15).*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Insert Table 2 here

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Recovering Quality of Life 20-item version (n=22)**

**Mental health.**

Table 3 shows that mental health scores improved significantly across admission [*F*(2,42) = 5.42, *p* < .01] with a large effect size ($η^{2}$**p = .21). Post hoc pairwise comparisons showed that the increase in scores was significant between admission and 12 months only (*p* < .05). The difference in scores within the first six months and the latter six months of admission did not reach significance.**

**Physical health.**

**Scores on the physical health item did not significantly change over time** [*F*(2,42) = 1.53, *p* = .23], **indicating that admission to the DBT IPU did not lead to an improvement in patients’ self-reported physical health.**

**Clinical Outcomes in Routine Evaluation – Outcome Measure (n=16)**

**Global distress.**

**The reduction in global distress scores was significant across timepoints [*F*(2,28) = 13.28, *p* < .001], with a large effect size (**$η^{2}$**p= .47). Pairwise comparison showed significant change at both 6 months (*p* < .05) and 12 months (*p* < .01), in comparison to baseline. The change in scores between 6 and 12 months post-admission was not significant.**

**Wellbeing.**

**Scores on the wellbeing subscale reduced significantly across DBT admission [*F*(2,30) = 6.64, *p* < .01], with a large effect size (**$η^{2}$**p= .30). Post hoc pairwise comparisons revealed that, in comparison to baseline, scores at both 6 and 12 months were significantly lower (both *p* < .05). The difference in scores at 6 and 12 months was non-significant.**

**Problems.**

**The reduction in scores on the problem subscale across admission was significant [*F*(2,30) = 10.36, *p* < .001], with a large effect size (**$η^{2}$**p=.41). Post hoc pairwise comparisons showed that scores at 6 months and 12 months were significantly lower compared to baseline (*p* < .05 and *p* < .001, respectively). The change in scores between 6 and 12 months was not significant.**

**Functioning.**

**The reduction in functioning scores was significant across admission [*F*(2,30) = 11.67, *p* < .001], with a large effect size (**$η^{2}$**p= .44). Post hoc pairwise comparisons showed that, compared to baseline, the change in scores was significant at both 6 months and 12 months post-admission (*p* < .05 and *p* < .01, respectively).**

**Risk.**

**The reduction in risk scores across admission reached significance [*F(*2,30) = 10.39, *p* < .001], with a large effect size (**$η^{2}$**p= .41). Pairwise comparisons showed that this reduction was significant at 12-months post-admission (*p* < .01), when compared to baseline scores. A non-significant trend for lower scores at 6-months compared to baseline emerged.**

**Health of the Nation Outcomes Scales for Users of Secure and Forensic Services (n=36)**

**Severe disturbance.**

**Results of the non-parametric Friedman test showed that severe disturbance scores significantly reduced across admission [**χ2(2)=19.75, *p* <.001]. Post hoc comparisons with Wilcoxon signed-rank tests revealed that**, compared to scores at baseline, severe disturbance rates were significantly lower at 6 (z = -3.75, *p* < .001, *r* = .44) and 12 months (z = -3.20, *p* < .01, *r* = .38), both results with a medium effect size. Scores at 6 and 12 months were comparable (z = -.24, *p* = .81).**

**Personal wellbeing.**

**Results of the non-parametric Friedman test showed that scores of personal wellbeing did not significantly change during admission [***χ*2(2) = 2.36, *p* =.31]**.**

**Emotional wellbeing.**

**Emotional wellbeing scores significantly reduced across admission [***χ*2(2) = 8.96, *p* < .05]. Post hoc Wilcoxon signed-rank tests showed that, in comparison to baseline, scores were significantly lower at 6 (z = -3.16, *p* < .01, *r* = .37) and 12 months (z = -3.03, *p* < .01, *r* = .36). Scores at 6 and 12 months were comparable (z = -.27, *p* = .79).

**Socioeconomic status.**

**Socioeconomic status (SES) scores changed significantly across admission [***χ2*(2) = 7.72, *p* < .05]**. Nevertheless, the** post hoc comparisons with Wilcoxon signed-rank tests were not significant (T0 -T1: z = -1.89, *p* = .06; T0-T2: z = -1.06, *p* = .29; T1-T2: z = -.94, *p* = .35).

**Risk.**

**The decrease in risk scores across admission was significant [***χ*2(2) = 17.40, *p* <.001]. Post-hoc Wilcoxon signed-rank tests showed that, in comparison to baseline, scores significantly reduced at 6 (z = -2.52, *p* = .01, *r* = .30) and 12 months (z = -3.24, *p* = .001, *r* = .38). The difference between 6 and 12 months was also significant (z = -2.41, *p* < .01, *r* = .28).

**Need.**

**The decrease in need scores was significant across admission [***χ2*(2) = 12.44, *p* =.01]. **Post-hoc Wilcoxon signed-rank tests showed that scores at 12 months significantly reduced when compared to baseline (z = -2.01, *p* = .012, *r* = .24). The difference in scores between baseline and 6 months (z = -2.01, *p* = .04) and between 6 months and 12 months (z = -1.99, *p* = .04) was not significant.**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Insert Table 3 here

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Risk Behaviours and Restrictive Practices (n=41)**

The average number of incidents of risk behaviours (self-harm, violence against objects, violence against people) and restrictive practices (restraints, seclusions) reduced across the 12-month admission period (see Table 4).

**Self-harm.**

The reduction in incidents of self-harm across admission was significant **[***χ*2(2) = 17.12, *p* <.001]. Post hoc Wilcoxon signed-rank tests showed that the reduction in incidents of self-harm was significant at 6 (z = -2.76, *p* < .01, *r* = .30) and 12 months (z = -3.16, *p* < .01, *r* = .35), when compared to baseline. The difference between 6 and 12 months was not significant (z = -1.91, *p* = .06).

**Violence.**

The reduction in incidents of violence against other people and/or objects approached significance [*χ2*(2) = 5.90, *p* =.052]. Post hoc Wilcoxon signed-rank tests showed that the difference in incidents of violence was significant between baseline and 6 months only (z = -2.85, *p* < .01, *r* = .45). No significant differences were apparent between 6 and 12 months (z = -.31, *p* = .75) nor baseline and 12 months (z = -2.18, *p* = .03).

**Aggregated aggression score (AAS).**

The mean AAS, weighted by type and severity of risk behaviour, significantly changed over time **[***χ2*(2) = 14.97, *p* <.01]. Results of the Wilcoxon signed-rank test show that the AAS significantly reduced at 12 months, compared to baseline (z = -3.33, *p* < .001, *r* = .37). The change between baseline and 6 months, and 6 and 12 months were not significant (z = -1.79, *p* = .07; and z = -2.16, *p* = .03, respectively).

**Restraint and seclusion.**

The use of restraints significantly reduced across admission [*χ2*(2) = 13.65, *p* <.01]. Post hoc Wilcoxon signed-rank tests showed that the difference in the use of restraints was significant between baseline and 12 months (z = -3.29, *p* < .01, *r* = .36). Restraint episodes did not differ between baseline and 6 months (z = -2.32, *p* = .02) nor between 6 and 12 months (z = -1.19, *p* = .23).

The use of restraints also significantly reduced across admission [*χ2*(2) = 10.38, *p* <.01]. Post hoc Wilcoxon signed-rank tests showed that the change in rates of seclusion was significant between baseline and 6 months (z = -2.72, *p* < .01, *r* = .30). The change in rates of seclusion between baseline and 12-months approached significance (z = -2.40, *p* = .016). Rates of seclusion between 6 and 12 months were comparable (z = -.17, *p* = .86).

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Insert Table 4 here

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 **Discussion**

The current study aimed to evaluate changes in outcomes for women with EUPD admitted to a specialist DBT service, across a 12-month admission period. The study built upon an earlier evaluation (Fox et al., 2014) of the comprehensive DBT programme being delivered within the service, prior to the adoption of a value-based model of care. The findings of the current study highlighted significant improvements across a number of patient- and clinician-reported outcomes, as well as risk behaviours and restrictive practices, over a 12-month period.

Overall, admission to the DBT IPU led to significant improvements in a number of outcomes relating to quality of life (ReQOL), psychological functioning (CORE-OM), health and social functioning, and security risks and needs (HoNOS-secure). The findings also highlighted behavioural changes, specifically improvements in the frequency of incidents of self-harm and reductions in the use of restrictive practices. Whilst reductions in incidents of violence against persons and objects bordered significance, this may be explained by the fact that the frequency of these incidents at baseline was already low; the findings still support a trend towards the reduction in their frequency across admission, as has been evidenced in other evaluations of DBT in secure psychiatric settings (e.g. Evershed et al., 2003). When accounting for the type and severity of incidents, analysis showed a significant reduction in AAS scores over the 12-month admission period.

Generally, the most significant improvements in outcomes occurred over the 12-month admission period, suggesting that a short-term 6-month admission period is not substantial in eliciting significant improvements. Exceptions were apparent with significant reductions in the use of seclusions and rates of incidents of externally directed violence within the initial 6-month admission period. Improvements on the ‘severe disturbance’ HoNOS subscale were slightly more significant in the initial 6 months of admission than the changes across the 12-month admission period, although this discrepancy was minimal. The reductions shown in risk behaviours and the use of restrictive practices across admission also mirrored changes in risk ratings on the HoNOS-secure; the need for a secure environment, as rated by staff, reduced significantly in the latter six-month period of admission, following the significant reduction in risk behaviours and need for seclusion as evidenced by patients within the first six months of admission.

Further exceptions were apparent, with no significant differences in measures of physical health (ReQoL), socioeconomic status and personal wellbeing (HoNOS) during admission. Nevertheless, there are a number of important caveats to consider when interpreting these findings. Firstly, physical health was measured by a singular item, as self-reported on the ReQoL (item 21), and thus substantiated conclusions around the effectiveness of the service in improving physical health outcomes cannot be drawn. Additionally, the non-significant change in ratings of socioeconomic status, a subscale comprised of the items ‘problems with living conditions’ and ‘problems with occupation and activities’, is likely a reflection of the setting itself; patients were detained to a secure inpatient service and thus were out of employment with restricted opportunities for activities.

Furthermore, whilst ‘personal wellbeing’ scores did not significantly change during admission, average scores on this subscale were somewhat low at each timepoint. Additionally, patient-reported ratings of ‘wellbeing’ on the CORE-OM did significantly change over admission, suggesting that patients perceived there to be significant improvements in this domain; given that the HoNOS-secure is a clinician-rated measure, this discrepancy in ratings for personal wellbeing may reflect divergence in clinicians’ and patients’ perceptions, which merits further exploration.

Despite these non-significant findings, the changes in outcomes reported here are largely consistent with those reported in the previous evaluation of the DBT service (Fox et al., 2014) which too evidenced significant improvements in outcomes pertaining to symptoms, functioning and security needs as well as risk behaviours, with moderate to large effect sizes. There is some discrepancy in the point of admission for which the greatest improvements in outcomes occurred, however. Whereas Fox et al. (2014) reported the most significant improvements in outcomes between baseline and 6 months post-admission, there was a tendency for the greatest changes to occur over a longer duration of admission, between baseline and 12-months post-admission, within the current study. Given that the outcomes explored within the current study were selected through a process of patient consultation, this could suggest that outcomes which are important to patients take longer to manifest, whilst outcomes selected by clinicians, as utilised in the previous evaluation, are sensitive to changes occuring within a shorter time period. The exception to this trend was in the frequency of ‘seclusions’, where a significant reduction was found between baseline and 6 months only. This reduction in rates of seclusion within the first 6 months of admission mirrors the significant reductions in incidents of self-harm and violence that also occurred within the initial 6-month admission period.

**Limitations**

The current study reports on outcomes across 12-month admission to the DBT service. Whilst data was available at 18 months post-admission for some patients (n=18), inconsistency in the outcomes for which patient data was available precluded analysis at this timepoint. Nevertheless, of those who had been discharged from the service, the average length of stay exceeded one year. Further investigation into outcomes across a longer period of admission with a larger patient cohort is warranted to better understand the full outcomes achieved across admission to the DBT IPU. through to the point of discharge.

Furthermore, data was not available for the whole sample on all of the outcome measures assessed, and thus the findings reported here pertaining to the ReQoL and CORE-OM are reflective of three separate subsamples. It cannot be assumed that the significant improvements found in these outcomes apply across the whole sample.

Finally, given that outcomes were evaluated in an IPU setting, it cannot be concluded that the improvements shown were a direct result of the DBT programme itself. Patients admitted to the service receive holistic care through a multidisciplinary team to address all areas of need.

**Conclusions**

The current findings expand on the previous results highlighting positive outcomes from admission to a DBT service for women with EUPD, across a wide range of domains. The findings also support the adoption of a value-based and outcomes-driven approach to healthcare, highlighting the assessment of meaningful patient outcomes as an effective approach to monitor progress. The results are largely consistent with those reported in an earlier evaluation of the DBT service, prior to the transition to a value-based healthcare model (Fox et al., 2014). Given that the outcomes were patient-valued and selected through collaborative consultation, the findings are indicative of progress towards ‘recovery’, as defined by those accessing the service. Future research directions should advance on these foundations and explore the measurement of ‘Quality Adjusted Life Years’ (QALYS) to quantify the value of care and demonstrate sustainable clinical impact.

**References**

Alderman, N., Knight, C., Stewart, I., & Gayton, A. (2011). Measuring behavioural outcome in neurodisability. *British Journal of Neuroscience Nursing, 7*(6), 691-695. https://doi.org/10.12968/bjnn.2011.7.6.691

Beck, A. T., & Steer, R. A. (1987). *Beck Depression Inventory*. San Antonio: The Psychological Corporation Inc.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences (2nd edition).* Hillsdale, NJ: Lawrence Erlbaum Associates

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101. https://doi.org/10.1111/1467-8721.ep10768783

Dahlke, F., Lohaus, A., & Gutzmann, H. (1992). Reliability and clinical concepts underlying global judgements in dementia: Implications for clinical research. *Psychopharmacology Bulletin, 28*(4), 425-432.

Department of Health. (1983). *Mental Health Act 1983.* England: Her Majesty’s Stationary Office.

Department of Health. (2007). *Mental Health Act 2007.* London: The Stationary Office.

Dickens, G. L., Sugarman, P., & Walker, L. (2007). HoNOS-Secure: A reliable outcome measure for users of secure and forensic mental health services. *Journal of Forensic Psychiatry and Psychology, 18*(4), 507-514. https://doi.org/10.1080/14789940701492279

Donabedian, A. (1966). Evaluating the quality of medical care. *The Milbank Memorial Fund Quarterly, 44*(3), 166-206. https://doi.org/10.1111/j.1468-0009.2005.00397.x

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE-OM. *The British Journal of Psychiatry, 180*(1), 51-60. https://doi.org/10.1192/bjp.180.1.51

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health, 9*(3), 247-255. https://doi.org/10.1080/jmh.9.3.247.255

Evershed, S., Tennant, A., Boomer, D., Rees, A., Barkham, M., & Watson, A. (2003). Practice-based outcomes of dialectical behaviour therapy (DBT) targeting anger and violence, with male forensic patients: A pragmatic and non-contemporaneous comparison. *Criminal Behaviour and Mental Health, 13*(3), 198-213. https://doi.org/10.1002/cbm.542

Field, A. (2013) *Discovering Statistics Using IBM SPSS Statistics (4th edition).* London: SAGE.

Fox, E. (2018). Delivering DBT in an inpatient setting. In M. A. Swales (Ed.), *The Oxford handbook of dialectical behaviour therapy,* pp. 645-670.

Fox, E., Krawczyk, K., Staniford, J., & Dickens, G. L. (2014). A service evaluation of a 1-year dialectical behaviour therapy programme for women with borderline personality disorder in a low secure unit. *Behavioural and Cognitive Psychotherapy, 43*(6), 676-691.

Girardi, A., Snyman, P., Natarajan, M., & Griffiths, C. (2018). Older adults in secure mental health care: Health, social wellbeing and security needs measures with HoNOS-secure across different age groups. *Journal of Forensic Psychiatry and Psychology, 29*(1), 1-20. https://doi.org/10.1080/14789949.2018.1459787

Guy, W. (1976). *ECDEU Assessment Manual for Psychoparmacology*. Rockville, MD: US Department of Health, Education and Welfare Public Health Service Alcohol, Drug Abuse and Mental Health Administration, pp. 218-222.

Kay, S. R., Wolkenfield, F., & Murrill, L. M. (1988). Profiles of aggression among psychiatric patients: II. Covariates and predictors. *The Journal of Nervous and Mental Disease, 176*(9), 547-557. https://doi.org/10.1097/00005053-198809000-00008

Keetharuth, A., Bjorner, J., Barkham, M., Browne, J., Croudace, T., & Brazier, J. (2020). An item response theory of analysis of an item pool for the recovering quality of life (ReQoL) measure. *Quality of Life Research.* https://doi.org/10.1007/s11136-020-02622-2

Keetharuth, A., Brazier, J., Connell, J., Bjorner, J., Carlton, J., Taylor Buck, E., Ricketts, T., McKendrick, K., Browne, J., & Barkham, M. (2018). Recovering Quality of Life (ReQoL): A new genetic self-reported outcome measure for use with people experiencing mental health difficulties. *The British Journal of Psychiatry, 212*(1), 42-49. https://doi.org/10.1192/bjp.2017.10

Keetharuth, A., Brazier, J., Connell, J., Carlton, J., Taylor Buck, E., Ricketts, T., & Barkham, M. (2017). *Development and validation of the Recovering Quality of Life (ReQoL) outcome measures*. Policy Research Unit in Economic Evaluation of Health and Care Interventions. Universities of Sheffield and York. EEPRU Research Report 050.

Kørner, A., Lauritzen, L., Abelskov, K., Gulmann, N. C., Brodersen, A. M., Wedervang-Jensen, T., & Kjeldgaard, K. M. (2007). Rating scales for depression in the elderly: External and internal validity. *The Journal of Clinical Psychiatry, 68*(3), 384-389.

Leon, A. C., Shear, M. K., Klerman, G. L., Portera, L., Rosenbaum, J. F., & Goldenberg, I. (1993). A comparison of symptom determinants of patient and clinician global ratings in patients with panic disorder and depression. *Journal of Clinical Psychopharmacology, 13*(5), 327-331.

Linehan. M. M. (1993). *Cognitive-Behavioural Treatment of Borderline Personal Disorder.* New York: The Guilford Press.

Linehan, M. M., Armstrong, H. E., Suarez, A., Allmon, D., & Heard, H. L. (1991). Cognitive-behavioural treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry, 63*, 757-766. https://doi.org/10.1001/archpsyc.1991.01810360024003

Maddison, P., Marlee, A. L., Webb, D. E., Berry, A. D. I., & Whitelock, T. (2016). The subscale structure of the HoNOS in a forensic population: A framework for a currency and tariff outcome system in mental health. *International Journal of Forensic Mental Health, 15*(4), 362-368. <https://doi.org/10.1080/14999013.2016.1208307>

McCloskey, H. (2001). Evaluation of the CORE outcome measure in a therapeutic forensic setting. *British Journal of Forensic Practice, 3*(2), 22-28. <https://doi.org/10.1108/14636646200100011>

Moulden, H. M., Mamak, M., Chaimowitz, G. (2020). A preliminary evaluation of the effectiveness of dialectical behaviour therapy in a forensic psychiatric setting. *Criminal Behaviour and Mental Health, 30*, 141-150.

National Institute for Health and Care Excellence. (2009). *Borderline personality disorder: recognition and management [CG78].* Available at: https://www.nice.org.uk/guidance/cg78

NHS England and NHS Improvement. (2016). *Delivering the Five Year Forward View for Mental Health: Developing quality and outcome measures.* Available at: https://www.england.nhs.uk/mentalhealth/wp-content/uploads/sites/29/2016/02/mh-quality-outcome.pdf

Pirkis, J. E., Burgess, P. M., Kirk, P. K., Dodson, S., Coombs, T. J., & Williamson, M. K. (2005). A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures. *Health and Quality of Life Outcomes, 28*(3). https://doi.org/10.1186/1477-7527-3-76

Porter, M. E. (2010). What is value in health care? *The New England Journal of Medicine, 363*(26)*,* 2477-2481. https://doi.org/10.1056/NEJMp1011024

Porter, M. E., & Lee, T. H. (2013). The strategy that will fix health care. *Harvard Business Review, 91*(10), 50-70.

Storebø, O. J., Stoffers-Winterling, J., Völlm, B., Kongerslev, M. T., Mattivi, J. T., Jørgensen, M. S., Faltinsen, E., Todorovac, A., Sales, C., Callesen, H. E., Lieb, K., & Simonsen, E. (2020). Psychological therapies for people with borderline personality disorder (Review). *Cochrane Database of Systematic Reviews.* http://doi.org/10.1002/14651858.CD012955.pub2

Sugarman, P., & Walker, L. (2007). *Health of the Nation Outcome Scales for Users of Secure and Forensic Services (HoNOS-Secure).* Available at: <https://www.rcpsych.ac.uk/events/in-house-training/health-of-nation-outcome-scales>

Tiffin, P., Daniel, M., & Carey, J. (2011, June). *The ability of the HoNOS-secure version 2b to predict care-setting.* Paper presented at International Association of Forensic Mental Health Services. Barcelona, Spain.

Tomlinson, M. F. (2018). A theoretical and empirical review of dialectical behavior therapy within forensic psychiatric and correctional settings worldwide. *International Journal of Forensic Mental Health, 17*(1), 72-95. https://doi.org/10.1080/14999013.2017.1416003

Wallang, P., Kamath, S., Parshall, A., Saridar, T., & Shah, M. (2018). Implementation of outcomes-driven and value-based mental health care in the UK. *British Journal of Hospital Medicine, 79*(6), 322-327. https://doi.org/10.12968/hmed.2018.79.6.322

Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B. G., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS): Research and development. *British Journal of Psychiatry, 172*, 11-18. https://doi.org/10.1192/bjp.172.1.11

World Health Organisation. (1992). *International Classification of Diseases and Related Health Problems, 10th Revision (ICD-10).* Geneva: World Health Organisation.

World Health Organisation. (2019). *International Classification of Diseases for Mortality and Morbidity Statistics*, *11th revision (ICD-11)*. Geneva: World Health Organisation.

Zuvekas, S. H., & Cohen, J. W. (2016). Fee-for-service, while much maligned, remains the dominant payment method for physician visits. *Health Affairs, 35*(3), 411-414. https://doi.org/10.1377/hlthaff.2015.1291