

The Effectiveness of 6 Versus 12 Months of Dialectical Behavior Therapy for Borderline Personality Disorder: A Noninferiority Randomized Clinical Trial

Online Supplement

Table of Contents

eTable1. Missingness for Times of Assessment Where Data Collected Was Used to Determine Noninferiority/ Superiority	2
eTable2. Baseline Demographics and Clinical Characteristics (Intent to Treat Sample).....	3
eTable3. Estimated Marginal Means and Changes in Number of Psychotropic Medications and Non-study Psychosocial Treatments Baseline to Month 24	5
eFigure1. Estimated Marginal Means Over Time and Tests of Noninferiority for Total Self-Harm	6
eFigure2. Estimated Marginal Means Over Time and Tests of Noninferiority for Non-suicidal self-injury (NSSI)	7
eFigure3. Estimated Probability of Suicide Over Time and Tests of Noninferiority	8
eFigure4. Estimated Marginal Means Over Time and Tests of Noninferiority for General Psychopathology (SCL-90-R)	9
eFigure5. Estimated Marginal Means Over Time and Tests of Noninferiority for Depression (BDI-II).....	10
eFigure6. Estimated Marginal Means Over Time and Tests of Noninferiority for Borderline Symptoms (BSL-23)	11
eFigure7. Estimated Marginal Means Over Time and Tests of Noninferiority for Skills Uptake (DBT-WCCL)	12
eFigure8. Estimated Marginal Means Over Time and Tests of Noninferiority for Interpersonal Functioning (IIP-64).....	13
eFigure9. Estimated Marginal Means Over Time and Tests of Noninferiority for Anger Expression (STAXI-2 AXO)	14
eAppendix1. Procedure for Determining Noninferiority Margins	15
eAppendix2. Questionnaire Used to Obtain Noninferiority Margins for Normally Distributed Outcomes.....	16
eAppendix3. Questionnaire Used to Obtain Noninferiority Margins for Non-Normally Distributed Outcomes. ...	19
eAppendix4. Distribution of Noninferiority Margins Provided by Non-Affiliated Clinicians	21
eAppendix5. Determination of Recovery from Self-Harm	22

eTable1. Missingness for Times of Assessment Where Data Collected Was Used to Determine Noninferiority/Superiority

		Time of Assessment				
Condition	Sample	6 Months	9 Months	12 Months	15 Months	24 Months
DBT-12	ITT	15.00%	20.83%	25.00%	26.67%	35.00%
	Per Protocol	0.00%	0.00%	3.85%	6.41%	17.90%
DBT-6	ITT	21.70%	23.33%	22.50%	30.00%	40.80%
	Per Protocol	7.78%	13.33%	12.22%	18.89%	31.10%
Overall	ITT	18.30%	22.08%	23.70%	28.33%	37.90%
	Per Protocol	4.17%	7.14%	8.33%	13.10%	25.00%

Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard); ITT, Intent to treat sample.

eTable2. Baseline Demographics and Clinical Characteristics (Intent to Treat Sample)

	Site		Length of Treatment		Intent to Treat <i>n</i> = 240
	CAMH <i>n</i> = 160	SFU <i>n</i> = 80	12-Months <i>n</i> = 120	6-Months <i>n</i> = 120	
Mean (SD) Age	27.44 [7.84]	28.39 [10.11]	27.26 [8.67]	28.25 [8.64]	27.75 [8.65]
Female	128 [80]	62 [78]	95 [79]	95 [79]	190 [79]
Marital Status:					
Married	23 [14]	15 [19]	22 [18]	16 [13]	38 [16]
Separated, Divorced, Widowed	13 [8]	8 [10]	9 [8]	12 [10]	21 [9]
Never married	124 [78]	57 [71]	89 [74]	92 [77]	181 [75]
Education:					
High School or Less	43 [27]	19 [24]	29 [24]	33 [28]	62 [26]
Some Post-Secondary	54 [34]	24 [30]	41 [34]	37 [31]	78 [33]
Post-Secondary	63 [39]	37 [46]	50 [42]	50 [42]	100 [42]
Employed	62 [39]	33 [41]	40 [33]	55 [46]	95 [40]
Income ^a					
<\$15,000	93 [58]	41 [51]	72 [60]	62 [52]	134 [56]
Between \$15,000 and \$29,000	43 [27]	35 [44]	36 [30]	42 [35]	78 [33]
Between \$30,000 and \$49,000	16 [10]	3 [4]	6 [5]	13 [11]	19 [8]
Median (IQR) Lifetime Suicide Attempts	1.00 [4.00]	2.00 [4.00]	1.00 [4.00]	2.00 [5.00]	1.00 [5.00]
Mean (SD) GAF	48.76 [7.29]	52.26 [8.08]	50.10 [7.75]	49.76 [7.73]	49.93 [7.73]
Lifetime Comorbid Axis I disorders					
Major depressive disorder	130 [81]	62 [78]	93 [78]	99 [83]	192 [80]
Panic disorder	51 [32]	32 [40]	42 [35]	41 [34]	83 [35]
Post-traumatic stress disorder	83 [52]	35 [44]	58 [48]	60 [50]	118 [49]
Any anxiety disorder	135 [84]	67 [84]	96 [80]	106 [88]	202 [84]
Any substance use disorder	132 [83]	59 [74]	100 [83]	91 [76]	191 [80]
Any eating disorder ^b	85 [53]	28 [35]	54 [45]	59 [49]	113 [47]
Current Comorbid Axis I disorders					
Major depressive disorder	61 [38]	34 [43]	46 [38]	49 [41]	95 [40]
Panic disorder	43 [27]	28 [35]	35 [29]	36 [30]	71 [30]
Post-traumatic stress disorder	52 [33]	29 [36]	37 [31]	44 [37]	81 [34]

eTable2. Baseline Demographics and Clinical Characteristics (Intent to Treat Sample) (continued)

	Site		Length of Treatment		Intent to Treat <i>n</i> = 240
	CAMH <i>n</i> = 160	SFU <i>n</i> = 80	12-Months <i>n</i> = 120	6-Months <i>n</i> = 120	
Current Comorbid Axis I disorders (cont'd)					
Any anxiety disorder	126 [79]	64 [80]	91 [76]	99 [83]	190 [79]
Any substance use disorder ^c	75 [47]	21 [26]	49 [41]	47 [39]	96 [40]
Any eating disorder ^d	35 [22]	15 [19]	18 [15]	32 [27]	50 [21]
Axis II cluster A diagnosis ^e	4 [3]	21 [26]	15 [13]	10 [8]	25 [10]
Axis II cluster B diagnosis (excl.BPD) ^e	8 [5]	16 [20]	12 [10]	12 [10]	24 [10]
Axis II cluster C diagnosis ^{e,f}	42 [26]	32 [40]	29 [24]	45 [38]	74 [31]
Mean (SD) Axis I Current disorders	2.82 [1.55]	3.07 [2.11]	2.85 [1.81]	2.94 [1.69]	2.90 [1.75]
Mean (SD) Axis I Lifetime disorders	5.01 [2.20]	5.11 [2.73]	5.14 [2.29]	4.94 [2.46]	5.04 [2.38]
Mean (SD) Axis II disorders (excl.BPD) ^g	0.34 [0.53]	1.05 [1.19]	0.53 [0.87]	0.63 [0.88]	0.58 [0.87]

Values reported are n (%) unless otherwise noted.

Chi-square and Fisher-exact tests (and T-tests for variables reporting a mean) revealed no significant differences between sites or conditions, except as noted.

Abbreviations: BPD, Borderline Personality Disorder; CAMH, BPD Clinic at Center for Addiction and Mental Health, Toronto, ON; GAF, Global Assessment of Function; SFU, DBT Centre in Burnaby, BC.

^a Proportion of participants earning between 15 and 29K was significantly higher at SFU ($p = .023$). There were no significant differences between conditions following randomization.

^b The proportion of participants currently or previously diagnosed with any eating disorder was higher at CAMH ($p = .009$). There were no significant differences between conditions following randomization.

^c The proportion of participants currently diagnosed with a substance use disorder was significantly higher at CAMH ($p = .005$). There were no significant differences between conditions following randomization.

^d The proportion of participants currently diagnosed with any eating disorder was higher for those assigned to 6-months of treatment ($p = .038$). There were no significant differences between conditions following randomization.

^e The proportion of participants diagnosed with a comorbid Cluster A, B, or C Axis II disorders was significantly higher at SFU ($p < .001$, $p < .001$, $p = .038$, respectively). There were no significant differences between conditions following randomization in case of Cluster A or B Axis II disorders.

^f The proportion of participants diagnosed with a comorbid Cluster C Axis II disorder was significantly higher in the 6-month condition ($p = .036$).

^g Participants at SFU were diagnosed with significantly more comorbid AXIS II disorders ($p < .001$). There were no significant differences between conditions following randomization.

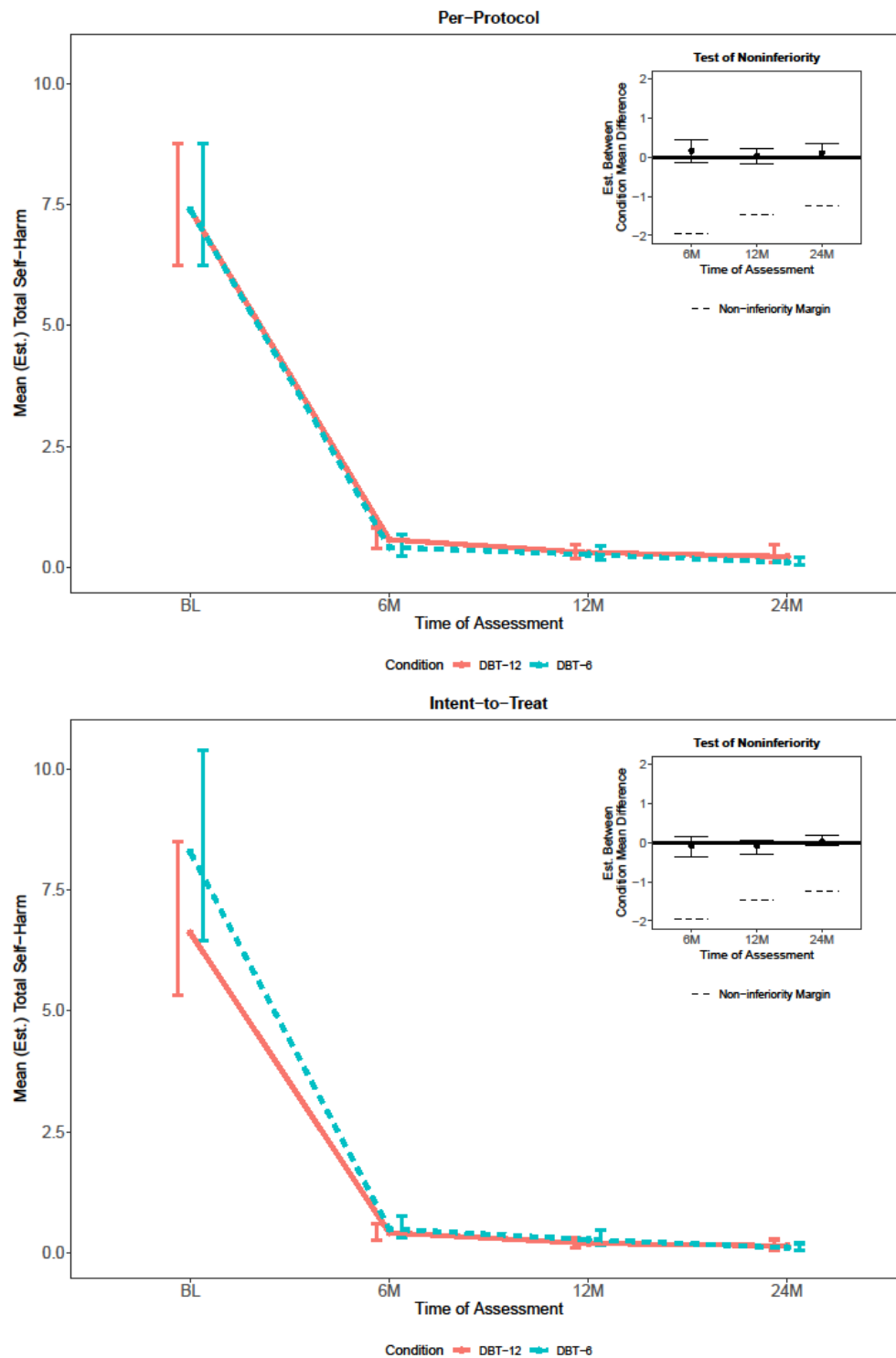
eTable3. Estimated Marginal Means and Changes in Number of Psychotropic Medications and Non-study Psychosocial Treatments Baseline to Month 24.

	Estimated Rate per 3 Months		Change (24M vs BL)		
	DBT-12 M [95% CI]	DBT-6 M [95% CI]	DBT-12 RR [95% CI]	DBT-6 RR [95% CI]	Ratio of Changes DBT-12 vs DBT-6 RR [95% CI]
Intent to Treat					
Psychotropic Medications					
Baseline	3.23 [2.72, 3.84]	2.94 [2.43, 3.57]	-	-	-
Month 24	2.19 [1.77, 2.73]	2.19 [1.74, 2.78]	0.68 [0.55, 0.83]	0.75 [0.61, 0.89]	0.91 [0.69, 1.19]
Psychosocial Treatments					
Baseline	1.41 [1.16, 1.68]	1.23 [1.00, 1.44]	-	-	-
Month 24	1.02 [0.74, 1.35]	1.01 [0.74, 1.28]	0.72 [0.52, 0.96]	0.82 [0.61, 1.03]	0.88 [0.60, 1.32]
Per Protocol					
Psychotropic Medications					
Baseline	3.21 [2.57, 3.91]	2.78 [2.13, 3.52]	-	-	-
Month 24	2.26 [1.72, 2.84]	2.12 [1.54, 2.77]	0.71 [0.55, 0.88]	0.76 [0.61, 0.92]	0.92 [0.68, 1.26]
Psychosocial Treatments					
Baseline	1.60 [1.21, 1.95]	1.29 [1.07, 1.55]	-	-	-
Month 24	1.14 [0.82, 1.55]	1.15 [0.86, 1.48]	0.71 [0.51, 1.00]	0.89 [0.69, 1.12]	0.80 [0.52, 1.21]

95% Confidence Intervals (95% CI) are bias-corrected bootstrapped, based on 1000 bootstrap samples.

Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard); RR, relative rate

eFigure1. Estimated Marginal Means Over Time and Tests of Noninferiority for Total Self-Harm

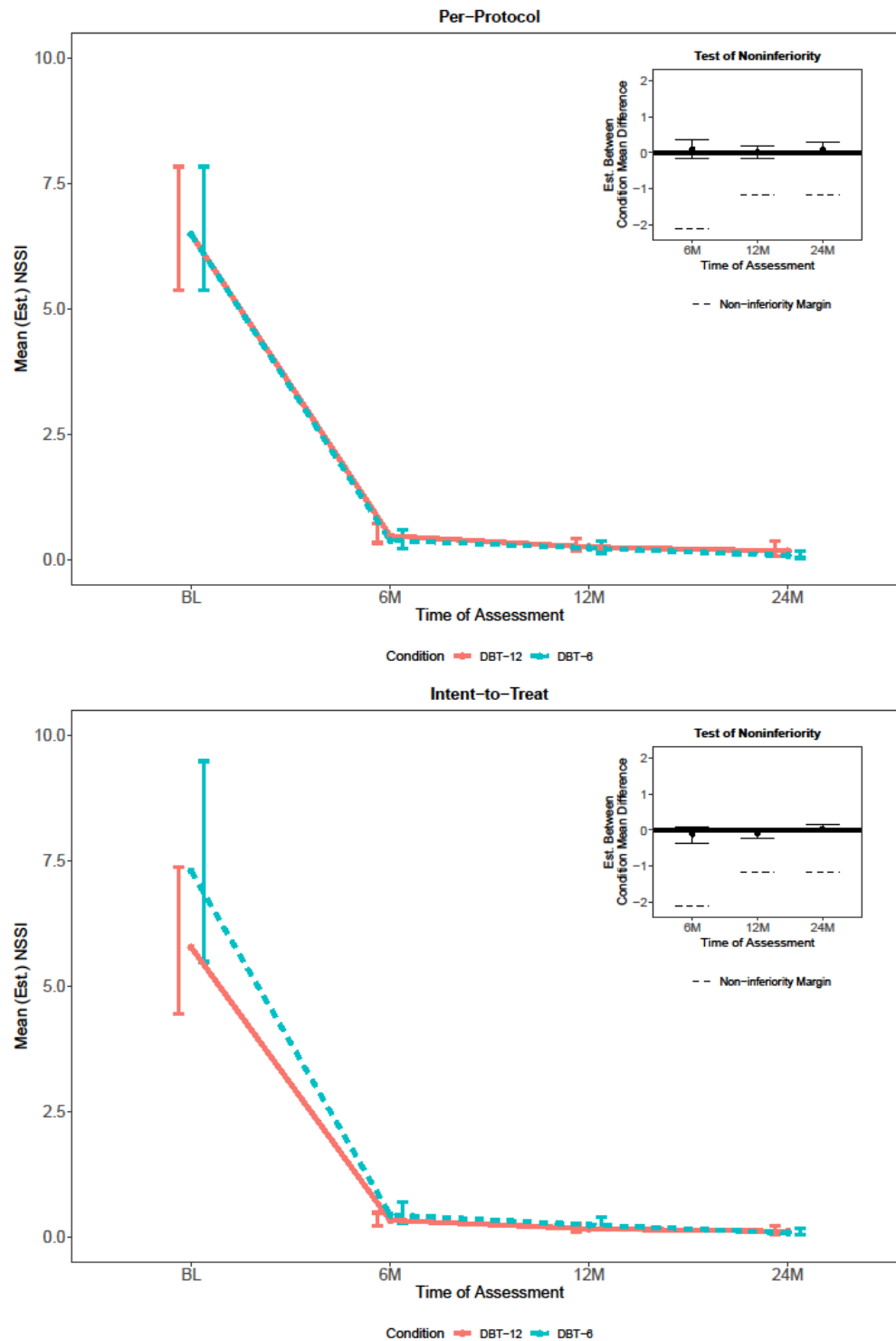


Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in total self-harm: 6M = -1.94, 12M = -1.47, 24M = -1.25.

As rates of total self-harm were measured over the three months previous to assessment, rates for 6 and 12-months are those assessed at the 9 and 15 months assessments and represent the rates over the 3 months immediately following end of 6 months and 12-months of treatment, respectively.

Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard).

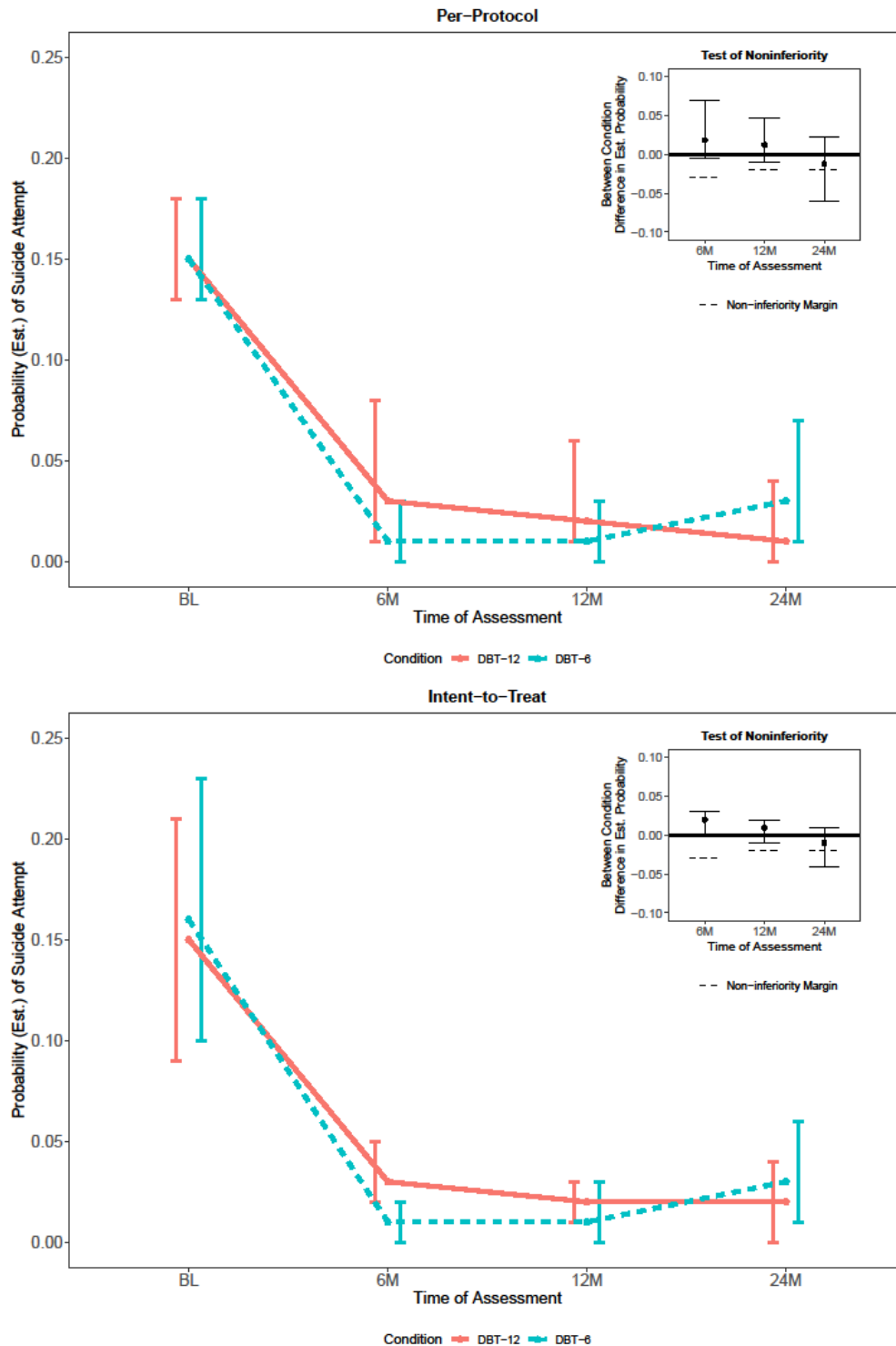
eFigure2. Estimated Marginal Means Over Time and Tests of Noninferiority for Non-suicidal self-injury (NSSI)



Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in non-suicidal self-injury: 6M = -2.12, 12M = -1.18, 24M = -1.17.

As rates of non-suicidal self-injury (NSSI) were measured over the three months previous to assessment, rates for 6 and 12-months are those assessed at the 9 and 15 months assessments and represent the rates over the 3 months immediately following end of 6 months and 12-months of treatment, respectively.

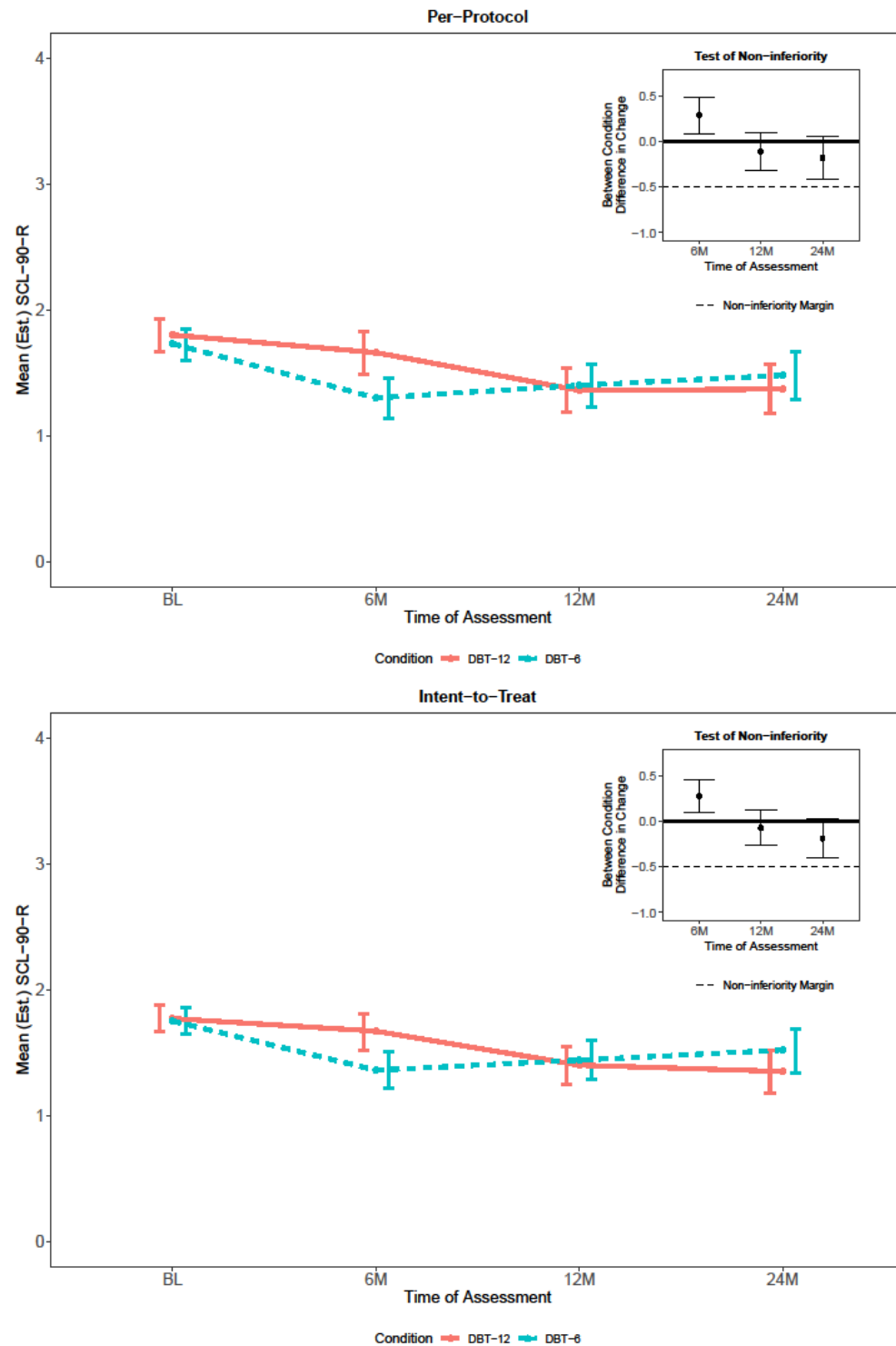
eFigure3. Estimated Probability of Suicide Over Time and Tests of Noninferiority



Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in probability of suicide: 6M = -.03, 12M = -.02, 24M = -.02.

As probability of suicide was measured over the three months previous to assessment, probabilities for 6 and 12-months is that assessed at the 9 and 15 months assessments and represents the probability of suicide over the 3 months immediately following end of 6 months and 12-months of treatment, respectively.

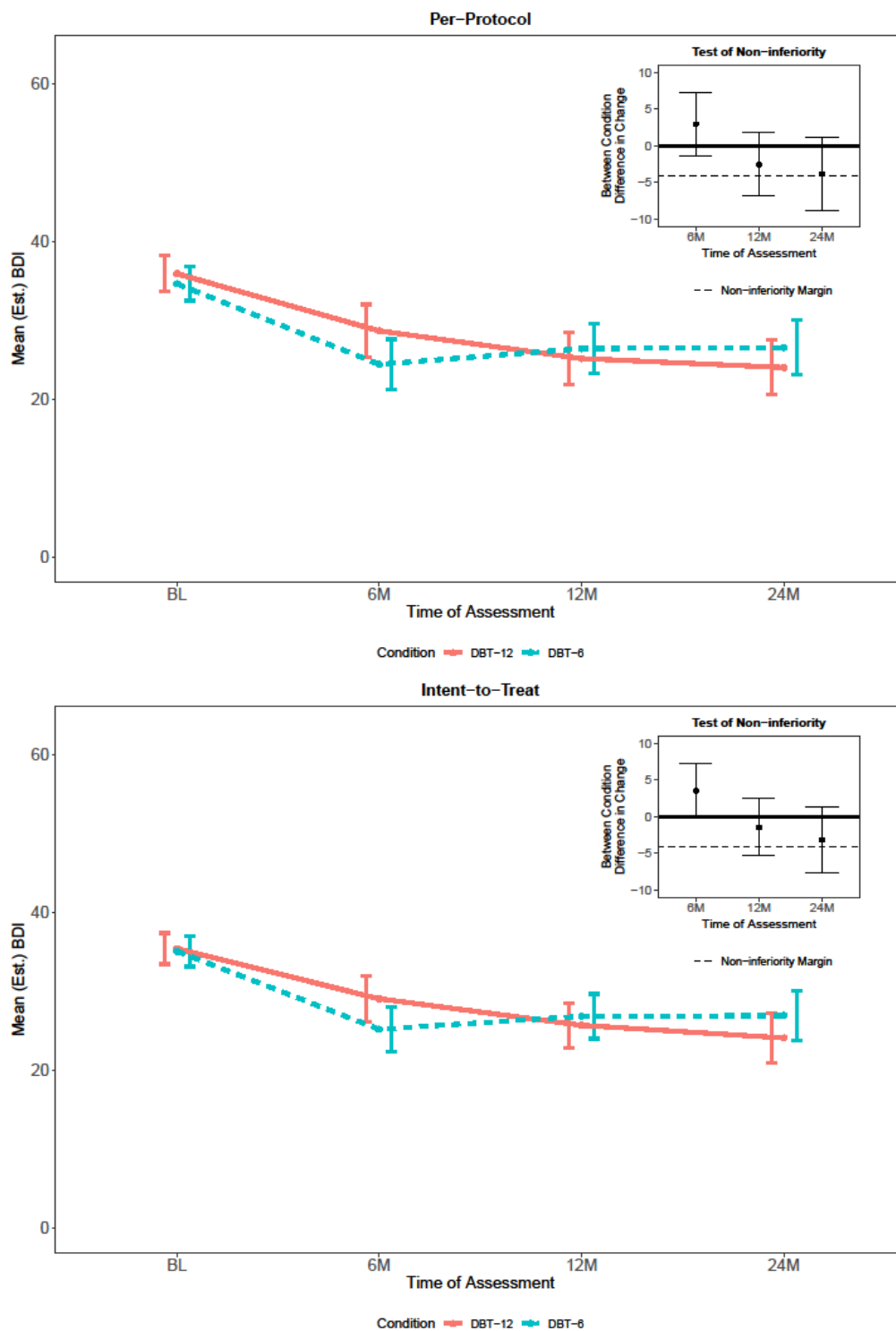
eFigure4. Estimated Marginal Means Over Time and Tests of Noninferiority for General Psychopathology (SCL-90-R)



Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in change in General Psychopathology (SCL-90-R): -.50.

Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard); SCL-90-R, Global Severity Index of the Symptom Checklist 90 Revised.

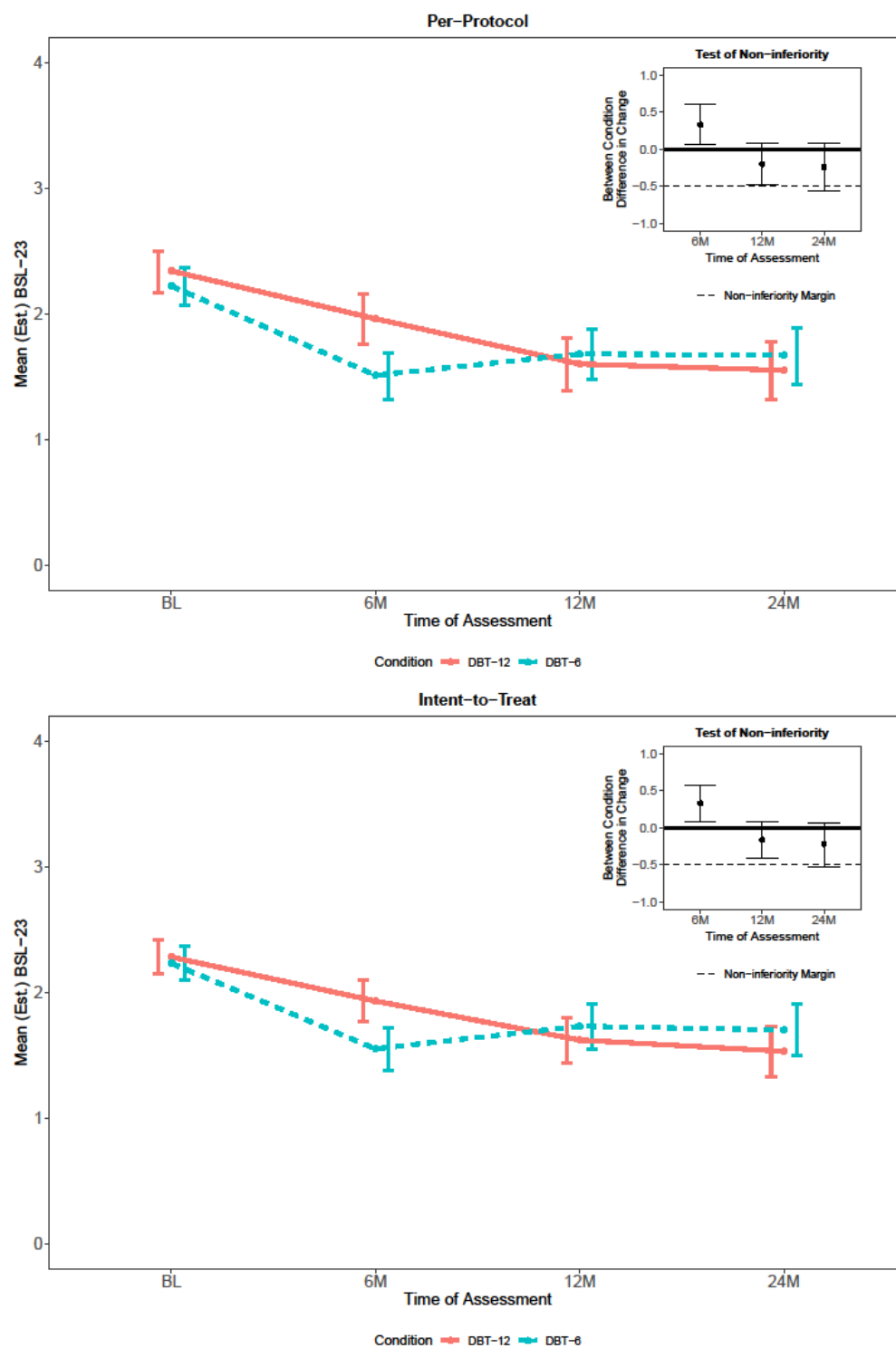
eFigure5. Estimated Marginal Means Over Time and Tests of Noninferiority for Depression (BDI-II)



Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in change in Depression (BDI-II): -.50.

Abbreviations: BDI, Beck Depression Inventory-II; DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard).

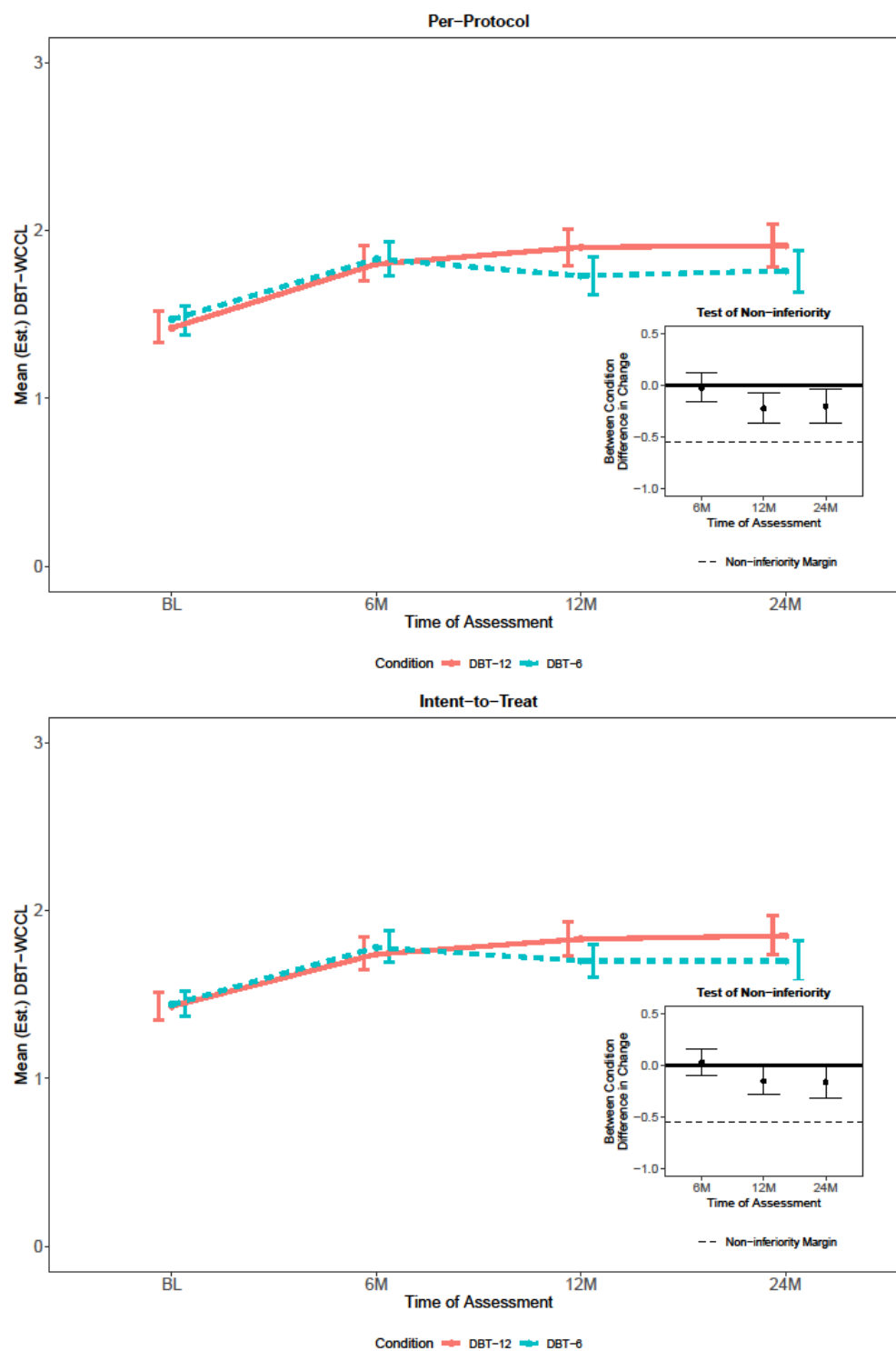
eFigure6. Estimated Marginal Means Over Time and Tests of Noninferiority for Borderline Symptoms (BSL-23)



Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in change in Borderline Symptoms (BSL-23): -.50.

Abbreviations: BSL-23, Borderline Symptom List-23; DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard).

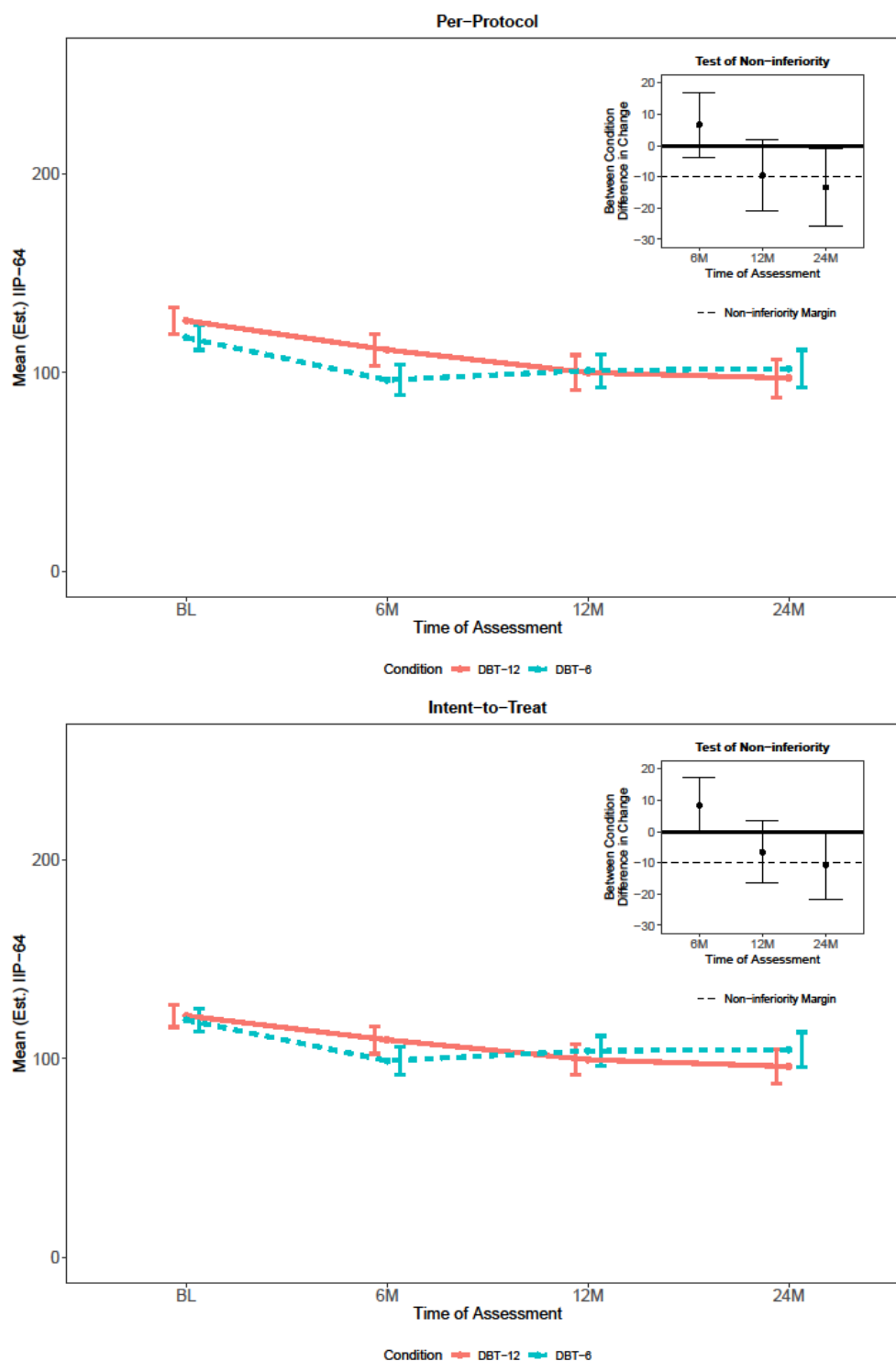
eFigure7. Estimated Marginal Means Over Time and Tests of Noninferiority for Skills Uptake (DBT-WCCL)



Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in change in Skills Uptake (DBT-WCCL): -.55

Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard); DBT-WCCL, Dialectical Behaviour Therapy Ways of Coping Checklist, Skills Use Subscale;

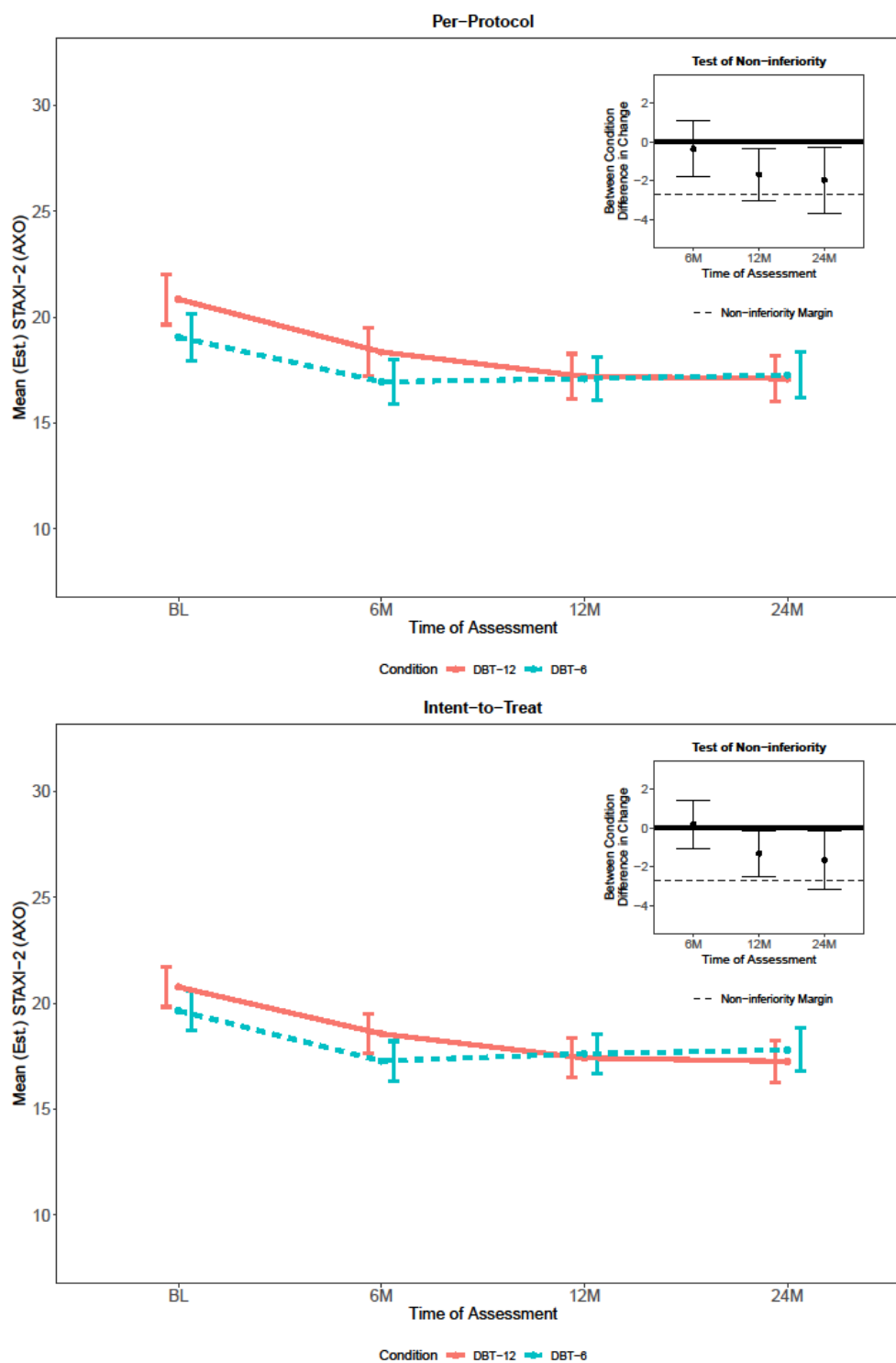
eFigure8. Estimated Marginal Means Over Time and Tests of Noninferiority for Interpersonal Functioning (IIP-64)



Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in change in Interpersonal Functioning (IIP-64): -10.00.

Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard); IIP-64, Inventory of Interpersonal Problems (64) Total Raw Score.

eFigure9. Estimated Marginal Means Over Time and Tests of Noninferiority for Anger Expression (STAXi-2 (AXO))



Noninferiority margin used to determine noninferiority of DBT-6 based on between condition difference in change in Anger Expression (STAXi_2 (AXO)): -2.67.

Abbreviations: DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration)

eAppendix1. Procedure for Determining Noninferiority Margins.

Noninferiority margins for all outcomes were set by a panel of clinicians expert in the treatment of BPD and, except as noted below, blinded to study results. Questionnaires were sent out in two waves.. In the first wave, clinicians were asked to provide the noninferiority margin for normally distributed secondary outcomes (e.g. Depression measured using the BDI-IIS). Here they were asked to indicate how much lower DBT-6 improvement in the outcome, measured using the appropriate scale , could become before the clinician elected to continue treating clients using the standard 12-month treatment, thus foregoing the potential benefits associated with the shorter, 6-month treatment. Clinician responses were then averaged, giving the noninferiority margins shown in Tables 3. Responses were only included in the average if the clinician was at least familiar with a number of items from the scale. eAppendix2 shows the questionnaire used to obtain these noninferiority margins.

Noninferiority margins for the primary outcomes total self-harm and NSSI, were obtained in a second wave. Here clinicians were first shown estimated rates /probabilities of occurrence of each outcome for DBT-12 at 9, 15, and 24-months. They were then asked how much higher the same rates for DBT-6 could become before the clinician elected to continue treating clients using the standard 12-month treatment, thus foregoing the potential benefits associated with the shorter, 6-month treatment. These differences were then averaged giving the noninferiority margins shown in Table 2. eAppendix3 shows the questionnaire used to obtain these noninferiority margins.

One of the clinicians involved completed training at our Toronto site and as part of this training conducted therapy with three of the participants assigned to DBT-6. This clinician performed no participant assessments upon which our results are based and had no access to any of the study data or results prior to participating in our survey.

eAppendix2. Questionnaire Used to Obtain Noninferiority Margins for Normally Distributed Outcomes.

CLINICALLY IMPORTANT DIFFERENCE

We are seeking expert opinion on clinically meaningful change for a set of clinical measures as part of our noninferiority study. This study compares DBT implemented over a shorter 6-month period to standard DBT that is administered over a period of 12 months. In order to complete our study we need to determine the point we will consider 6-months of DBT treatment to be at least as good as standard 12-months DBT. For our study, we are defining this value to be the LARGEST difference in change on a set of clinical measures between two treatment protocols that would NOT be considered clinically meaningful.

What is a noninferiority study

(Adapted from <https://senguptasresearchacademy.com/noninferiority-study-design-in-clinical-trials/>)

Newer treatments may be only marginally better than the established treatment in terms of the primary outcome. However, new treatments may have other practical advantages such as lower cost, better safety, lower frequency of administration and better compliance, which makes them attractive options to replace the treatment being currently used. Hence, we would at least expect that the effect of the new treatment is at least as good as the old treatment. This concept of “at least as good as” is the fundamental basis for noninferiority study designs.

The null hypothesis in superiority trials is that there is no difference between the treatments and you try to prove this hypothesis wrong and conclude that there is a difference i.e. one treatment is superior to the other. However, if you end up accepting the null hypothesis, concluding that there is no difference does not necessarily suggest equivalence of the two treatments. In a noninferiority study, the null hypothesis is the opposite, i.e. the newer treatment is inferior to the existing treatment. We try to prove this wrong and conclude that there is no difference between treatments i.e. they are non-inferior.

Please let us know your name before beginning:

SECTION 1: CLINICAL SCALES

Q1. How familiar are you with each of the following measures:

	0 No idea what the scale is	1 I have heard of the scale	2 I know what the scale is trying to measure	3 I know some of the specific questions on this	4 I am extremely familiar with this scale
a. BDI-II Beck Depression Inventory II	☐	☐	☐	☐	☐
b. BSL-23 Borderline Symptom List	☐	☐	☐	☐	☐
c. DBT-WCCL DBT-Ways of Coping Checklist	☐	☐	☐	☐	☐
d. IIP-64 Inventory of Interpersonal Problems	☐	☐	☐	☐	☐
f. SCL-90-R Symptoms Checklist 90 Revised (Global Severity Index; GSI)	☐	☐	☐	☐	☐
g. STAXI-2 State and Trait Anger Expression Inventory (anger expression subscore)	☐	☐	☐	☐	☐

SECTION 2: POINT CHANGE

When comparing two treatments, there is no clinically important difference in change between the two treatments when....

The change between pre- and post- scores for Treatment A differs from the change between pre- and post- scores for Treatment B by no more than ____ points *

Example

- For Treatment A, the average pre-test score was 46. The average post-test score was 30. The change in score XX for Treatment A was a decrease of 16 points.
- For Treatment B, the average pre-test score was 45. The average post-test score was 32. The change in score XX for Treatment B was a decrease of 13 points.
- The difference in changes for XX between the two protocols was therefore 3 points ($16-13=3$).

	Pre	Post	Point Difference
Treatment A (new)	46	30	16
Treatment B (standard)	45	32	13
			3

Is this 3 point difference in change in score XX between the two treatments NOT clinically important? What is the LARGEST difference in points that is NOT clinically significant?

Q2. What is the LARGEST difference in points that is NOT clinically significant?**a) BDI-II Beck Depression Inventory-II:**

total score 0-63 (lower is less depressed)

b) BSL-23 Borderline Symptom List:

mean score 0-4 (lower is less BPD symptoms)

c) DBT-WCCL DBT-Ways of Coping Checklist:

skills use subscale mean score 0-3 (higher is poorer coping)

d) d. IIP-64 Inventory of Interpersonal Problems:

total raw score 0-256 (higher is more interpersonal problems)

e) SCL-90-R Symptoms Checklist 90 Revised:

GSI score 0-4 (lower is less distress)

f) g. STAXI-2 State and Trait Anger Expression Inventory:

Anger expression subscore 8-32 (lower is less anger expression)

eAppendix3. Questionnaire Used to Obtain Noninferiority Margins for Non-Normally Distributed Outcomes.

Clinically Meaningful Difference Survey

Our FASTER DBT study compared 6- to 12-months DBT for 240 patients diagnosed with borderline personality disorder who recently engaged in chronic self-injurious behaviour (i.e. at least two episodes of non-suicidal self-injury or suicide attempts in the past 5 years, including at least one episode in the past 8 weeks).

We conducted a noninferiority trial in which we tested whether the newer 6-month DBT (i.e., shorter) is no worse than the standard 12-month DBT. To do so, we need to determine the point at which differences between the shorter 6-month DBT and the standard 12-month DBT on a set of outcomes are so great that one would continue treating patients using standard 12-month DBT, despite the obvious benefit of a shorter treatment (i.e., less time in treatment). We need to determine: what difference between 6-month and 12-month DBT in the mean number of total self-harms (total of suicide and NSSIs), NSSIs, and emergency room visits would be acceptable, and what difference between 6-month and 12-month DBT in the percentage of individuals attempting suicide and the proportion of individuals admitted to a psychiatric ward during the past 3-months would be acceptable. As an example, suppose you are comparing the mean number of total self-harms at the end of 12 months for 6-month DBT to that for standard 12-month DBT. You find patients treated using the 6-month DBT had 2 more episodes of self-harm than patients treated using standard 12-month DBT. Would you decide that it would be better for patients to do standard 12-month DBT on the basis of this finding? Or, would you think that 2 more episodes of self-harm would be worth the benefit of a shorter (6 months) treatment? In other words, a difference of 2 more episodes would not be high enough to suggest patients should do an additional 6 months of treatment. For all the comparisons, we are measuring the number of events (in this case self-harm) that occur in the 3 month period immediately following this comparison timepoint; so these extra 2 episodes of self-harm are occurring between 12 months and 15 months and hereafter we refer to as count at the **end of 12 months**.

Let's ask this question again, but this time give you some actual numbers from our study. At the beginning of the study, the mean number of total self-harms for all patients during the past 3 months was 7.39. By the end of 12 months, this mean number had dropped to .32 in the standard 12-month DBT group. How high would this mean number need to be for patients treated using 6-month DBT before you would decide that it would be better to treat patients using standard 12-month DBT?

If your answer is 2.32, then you would be saying that as long as the mean rate of total self-harms during the past three months for patients treated using 6-month DBT was 2.32 by the end of 12 months, you would consider treating patients using 6-month DBT. This would also imply that if the mean rate of total self-harms (during the past three months) was more than 2.32 among patients treated using 6-month DBT, you would use standard 12-month DBT to treat patients despite the obvious benefits associated with a shorter treatment.

Q_Name **Please let us know your name before beginning:**

Now, we are going to provide you with the mean number of total self-harms (suicide and NSSIs), NSSI's, and emergency room visits for patients treated using standard 12-month DBT. We would like you to indicate how high these mean numbers for each of these outcomes would have to be for patients treated using 6-month DBT before you would recommend 12 months of treatment (i.e., you would think the tradeoff of a shorter treatment would not be advantageous).

A. Total Self-Harms (NSSI's plus suicide attempts)

Assume the mean number of total self-harms at baseline for both groups was **7.39**.

A1) The mean number of total self-harms for the patients treated using standard 12-month DBT was **0.57 at the end of 6 months** of treatment (this is the end of treatment for the 6-month patients and mid-treatment for the 12-month patients). How high would the mean for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it)?

A2) The mean number of total self-harms for patients treated using standard 12-month DBT was **0.30 at the end of 12 months** (this is the end of treatment for the 12-month patients and 6 months post-treatment for the 6-month patients). How high would the mean for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it)?

A3) Finally, the mean number of total self-harms for patients treated using standard 12-month DBT was **0.22 at the end of 24 months** of treatment (this is the 1 year follow up for the 12-month patients and the 18 month follow up for the 6-month patients). How high would the mean for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it)

B. NSSIs (Non-suicidal self-injuries)

Assume the mean number of NSSIs at baseline for both groups was **6.48**.

B1) The mean number of NSSIs for the patients treated using standard 12-month DBT was **0.47 at the end of 6 months** of treatment (this is the end of treatment for the 6-month patients and mid-treatment for the 12-month patients). How high would the mean for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it)?

B2) The mean number of NSSIs for patients treated using standard 12-month DBT was **0.25 at the end of 12 months** (this is the end of treatment for the 12-month patients and 6 months post-treatment for the 6-month patients). How high would the mean for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it)

B3) Finally, the mean number of NSSIs for patients treated using standard 12-month DBT was **0.17 at the end of 24 months** of treatment (this is the 1 year follow up for the 12-month patients and the 18 month follow up for the 6-month patients). How high would the mean for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it).

Now, we are going to provide you with the percentage of patients treated using standard 12-month DBT attempting suicide (during the past 3 months) and the percentage of patients treated using standard 12-month DBT admitted to a psychiatric ward (during the past three months). We would like you to indicate how high the percentage of patients treated using 6-month DBT who attempted suicide would have to be before you would recommend 12-months of treatment (i.e., you would think the tradeoff of a shorter treatment would not be advantageous). how high the percentage of patients treated using 6-month DBT who were admitted to a psychiatric ward would have to be before you would recommend 12-months of treatment (i.e., you would think the tradeoff of a shorter treatment would not be advantageous).

C. Percentage of patients attempting suicide (in the previous 3 months)

Assume the percentage of patients attempting suicide (in the previous 3 months) at baseline for both groups was **19.6%**

D1) The percentage of patients treated using standard 12-month DBT who attempted suicide during the past 3 months was **3.03% at the end of 6 months** of treatment (this is the end of treatment for the 6-month patients and mid-treatment for the 12-month patients). How high would the percent for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it)?

D2) The percentage of patients treated using standard 12-month DBT who attempted suicide during the past 3 months was **2.52% at the end of 12 months** (this is the end of treatment for the 12-month patients and 6 months post-treatment for the 6-month patients). How high would the percent for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it)

D3) Finally, the percentage of patients treated using standard 12-month DBT who attempted suicide during the past 3 months was **1.58% at the end of 24 months** of treatment (this is the 12 month follow up for the 12-month patients and the 18 month follow up for the 6-month patients). How high would the percent for the 6-month patients have to be for you to recommend 12 months of treatment (i.e., no longer think the benefits of a shorter treatment are worth it)

eAppendix4. Distribution of Noninferiority Margins Provided by Non-Affiliated Clinicians

		Mean ^a	Trim Mean ^b	Median	Min	Max	Number responding
Non-normally distributed outcomes (differences, DBT-12 minus DBT-6)							
Total Self-Harm	6M	-1.94	-1.76	-1.43	-1.43	-3.00	5
	12M	-1.47	-1.63	-1.70	-0.20	-2.27	5
	24M	-1.25	-1.28	-1.28	-0.08	-2.35	5
Non-Suicidal Self-Injury (NSSI)	6M	-2.12	-2.02	-1.53	-1.53	3.03	5
	12M	-1.18	-1.07	-0.65	-0.05	2.65	5
	24M	-1.17	-0.99	-0.63	-0.08	2.83	5
Suicide ^c	6M	-0.03	-0.03	-0.03	0.00	-0.07	5
	12M	-0.02	-0.02	-0.02	0.00	-0.03	5
	24M	-0.02	-0.02	-0.02	0.00	-0.03	5
Emergency Department (ED) Visits	6M	-0.73	-0.21	-0.23	-0.03	-3.00	5
	12M	-0.51	-0.18	-0.13	-0.03	-2.00	5
	24M	-0.52	-0.20	(-0.27	-0.03	-2.00	5
Psychiatric Ward Admissions ^c	6M	-0.04	-0.04	-0.04	-0.03	-0.06	5
	12M	-0.04	-0.04	-0.04	-0.02	-0.05	5
	24M	-0.03	-0.03	-0.03	-0.02	-0.05	5
Normally distributed outcomes (between condition differences in change, DBT-12 minus DBT-6)							
General Psychopathology (SCL-90R)	All Times	-0.50	-0.43	-0.53	-0.20	-1.00	6
Depression (BDI-II)	All Times	-4.14	-4.20	-3.00	-2.00	-7.00	7
BPD Symptoms (BSL-23)	All Times	-0.50	-0.45	-0.40	-0.20	-1.00	6
Skills Uptake (DBT-WCCL)	All Times	-0.55	-0.50	-0.50	-0.20	-1.00	4
Interpersonal Functioning (IIP-64)	All Times	-10.00	-	-10.00	-5.00	-15.00	2
Anger Expression (STAXI-2 (AXO))	All Times	-2.67	-2.00	-2.50	-2.00	-4.00	4

Abbreviations: BDI-II, Beck Depression Inventory-II ; BSL-23, Borderline Symptom List-23; DBT-6, Dialectical Behavior Therapy (6 month duration); DBT-12, Dialectical Behavior Therapy (12 month duration – standard); DBT-WCCL (Skills Use), Dialectical Behaviour Therapy Ways of Coping Checklist, Skills Use Subscale; IIP-64, Inventory of Interpersonal Problems (64) Total Raw Score, Margin, margin used to determine noninferiority of DBT-6 based on difference in change in measure; SCL-90-R, Global Severity Index of the Symptom Checklist 90 Revised; STAXI -2(AXO), State-Trait Anger Expression Inventory-2, Anger Expression Out sub-scale.

^aMean of responses used as final noninferiority margin as the mean best captured the potential range of how clinicians might define clinical significance in the wider community of clinicians.

^bTrim mean calculated by removing highest and lowest margin supplied, except for skills uptake, where only two margin were supplied.

^c Results for suicide and psychiatric ward admissions represent the probability of a reported suicide attempt and the probability of admission to a psychiatric ward, respectively.

eAppendix 5. Determination of Recovery of Self-Harm

Participants completing treatment (and providing complete data on self-harming behavior to the end of the specific reporting period) were first classified into one of three groups based on severity of self-harms over the three months prior to baseline:

- One or more self-harms including a suicide (Level 3 severity),
- multiple self-harms, but no suicide attempts (Level 2 severity),
- one self-harm but no suicide attempts (Level 1 severity).

Participants were then classified into groups based on self-harms reported at the end of treatment (i.e. over the three months immediately following 6 months of treatment for DBT-6, and 12 months of treatment for DBT-12) as follows:

- Did they report any self-harms and, if so, into additional groups based on severity ratings used at baseline.
- Finally, participants were finally classified into one of four categories based on changes in self-harming categories baseline to end of treatment:

Deteriorated, level of severity of self-harming behavior increased;

No change, no change in level of severity of self-harming behavior;

Partial Recovery, self-harming behavior still evident, but severity decreased;

full recovery, no evidence of self-harming behavior.

The same process was used to classify participants into groups based on changes in self-harm severity ratings at the end of the followup period (i.e. Over the three months immediately preceding study end at 24 months).