

**Daily and weekly mood ratings using a remote capture method in high-risk offspring of bipolar parents: Compliance and symptom monitoring**

Authors: Duffy, A.<sup>1,2</sup>, Keown-Stoneman, C.<sup>2,3</sup>, Goodday, S.M.<sup>2,4</sup>, Saunders, K.<sup>5</sup>, Horrocks, J.<sup>6</sup>, Grof, P.<sup>2</sup>, Weir, A.<sup>2</sup>, Hinds, C.<sup>5,7</sup> and Geddes, J.<sup>5</sup>

1. Department of Psychiatry, Queen's University, Kingston, ON, Canada
2. Mood Disorders Centre of Ottawa, Ottawa, ON, Canada
3. Applied Health Research Centre (AHRC), Li Ka Shing Knowledge Institute of St. Michael's Hospital, Toronto, Ontario, Canada
4. Department of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
5. Department of Psychiatry, University of Oxford, Oxford, UK
6. Department of Mathematics and Statistics, University of Guelph, Guelph, ON, Canada
7. Big Data Institute, University of Oxford, UK

Corresponding Author:

Anne Duffy, Department of Psychiatry  
Queen's University Student Wellness Services  
LaSalle Building, 146 Stuart Street  
Kingston Ontario, Canada, K7L 3N6  
Tel: 613-533-2508  
Fax: 613-533-2347

Email: [anne.duffy@queensu.ca](mailto:anne.duffy@queensu.ca)

Word Count:

3825

Declaration of interest

All authors have no conflicts to declare in relation to this manuscript

## High frequency mood ratings in high-risk youth

### **Abstract**

#### Objectives

To determine the compliance and clinical utility of weekly and daily mood symptom monitoring in adolescents and young adults at risk for mood disorder.

#### Methods

Fifty emerging adult offspring of bipolar parents were recruited from the Flourish Canadian high-risk cohort study along with 108 university student controls. Participants were assessed by KSADS/SADS-L semi-structured interviews and used a remote capture method to complete weekly and daily mood symptom ratings using validated scales for 90 consecutive days. Hazard models and generalized estimating equations were used to determine differences in summary scores and regularity of ratings.

#### Results

78% and 77% of high-risk offspring and 97% and 93% of controls completed the first 30 days of weekly and daily ratings, respectively. There were no differences in drop-out rates between groups over 90 days (high-risk  $p=0.2149$ ; controls  $p=0.9792$ ). There were no differences in mean summary scores or regularity of weekly anxiety, depressive or hypomanic symptom ratings between high-risk and control groups. However, high-risk offspring compared to controls had daily ratings indicating lower positive affect and higher negative affect ( $p=0.0317$ ). High-risk offspring with remitted mood disorder compared to those without had more irregularity in weekly anxiety and depressive symptom ratings and daily ratings of lower positive affect, higher negative affect, and higher shame and self-doubt ( $p=0.0365$ ).

#### Conclusions

Findings support that high-resolution symptom tracking may be a feasible and clinically useful approach to monitoring emerging psychopathology in young people at high-risk of mood disorder onset or recurrence.

#### **Key Words**

Bipolar disorder, high-risk, offspring, self-report symptoms, affect, self-esteem, remote capture, online, regularity, symptom monitoring, compliance

## High frequency mood ratings in high-risk youth

### Introduction

Prospective studies of weekly ratings using remote capture methods in adults with an established diagnosis of bipolar disorder have demonstrated good compliance and acceptability.<sup>1, 2</sup> Further, these studies have demonstrated a high prevalence of residual symptoms during remission<sup>1, 3</sup> and an increase in irregularity of symptoms prior to mood episode recurrences of either polarity.<sup>4</sup> To our knowledge, remote capture methods have not been systematically studied in high-risk populations. Yet, remote capture symptom ratings could serve as a low cost and effective method of monitoring the emergence of new or recurrent clinically important symptoms and advancing the identification of “ultra-high risk” individuals who would benefit from more intensive clinical surveillance.

Given the high estimated heritability, offspring of bipolar parents are an identifiable high-risk population ideally suited for prospective symptom monitoring and a population who might benefit from such clinical surveillance. To wit, a recent paper from the BIOS high-risk offspring study identified that anxiety and depressive symptoms (reported by parent) and mood lability (reported by offspring) at one time point (at baseline) predicted the onset of bipolar disorder, in part through an association with more proximal hypomanic symptoms.<sup>5</sup> Findings from the Canadian high-risk offspring study provided further evidence that clinician-determined hypomanic symptoms predicted a 5-fold increased risk of recurrent mood disorder.<sup>6</sup>

Most studies of symptom level psychopathology in high-risk offspring have relied on retrospective recall or assessment at one time point. Even in prospective studies, the subject is typically required to recall their mood over a fixed period of time (i.e.  $\geq 12$  months), which is subject to recall bias; for example, through the influence of current mood state, the most recent intense experience or the level of instability in symptoms on recall.<sup>7</sup> Clearly, more precise and sensitive measures of symptom level data are needed to determine the timing and strength of any associations with mood disorder onset and recurrence with greater accuracy.

This paper describes a consecutive 90-day collection of self-report mood symptom data using methods validated in the AMoSS prospective deep phenotyping study<sup>8</sup> in participants from the

## High frequency mood ratings in high-risk youth

Canadian Flourish high-risk offspring study.<sup>9</sup> Our objectives were to determine (i) compliance rates with these methods and (ii) the potential clinical utility of symptom monitoring in high-risk individuals. Specifically we hypothesized that (i) there would be comparable high rates of completed data over the first 30 and 90 day period of study using remote capture methods in high-risk and control groups, and (ii) symptom ratings would be higher and show more irregularity in high-risk offspring compared to controls, and in clinically well-remitted high-risk offspring with compared to those without a prior mood diagnosis - supporting the clinical utility of symptom monitoring using these methods.

## Subjects and Methods

### *Bipolar families*

Research families were identified through an adult bipolar patient attending a specialty mood disorder program in Ottawa and participating in other genetic and neurobiological research.<sup>10</sup> The DSM-IV diagnosis of BD I or II was based on SADS-L interviews conducted by a research psychiatrist and confirmed on blind consensus review using best estimate procedures taking into account all available research and clinical data.<sup>9</sup> Over the years, we expanded recruitment to include other similarly assessed adult siblings of the original BD probands who themselves met lifetime criteria for bipolar spectrum disorder (BDI, II, NOS, recurrent major depressive disorder). The families have remained actively involved in research and the original probands and their relatives have been repeatedly assessed for up to 40 years, hence confirming the stability of the diagnosis and therefore the risk status of the offspring.<sup>11-13</sup>

### *High-risk offspring*

As detailed previously,<sup>9</sup> enrolment in the Canadian Flourish high-risk offspring study included all assenting/consenting children from identified families who met the following inclusion criteria: (i) one parent meeting DSM-IV criteria for BD or a first degree relative with BD spectrum disorder as described above and (ii) age 5-25 years at baseline. Exclusion criteria: (i) other biological parent meeting lifetime DSM criteria for schizophrenia, recurrent major depressive or bipolar disorder and /or (ii) any offspring with a serious neurological disorder, intellectual handicap, and/or unable to comply with study protocol. For this remote capture

## High frequency mood ratings in high-risk youth

study, we were funded to enroll 50 consenting clinically well (never ill or fully remitted) high-risk offspring who were at least 13 years of age or older. We therefore enrolled the first 50 eligible and consenting offspring at their regularly scheduled annual research visit.

### *Controls*

For the comparison group, we were funded to include the first 100 eligible and consenting undergraduate university students attending the University of Calgary. Participating students earned credits towards the mandatory research requirement for their program of study. Inclusion criteria: (i) well or in good quality clinical remission from a lifetime anxiety or mood disorder based on DSM-IV criteria. Exclusion criteria: (i) already meeting DSM-IV criteria for BD or psychotic disorder or (ii) having a parent with a lifetime diagnosis of BD or schizophrenia based on participant report.

### *Procedure*

All participants completed written informed consent forms approved by the local Ottawa Chesapeake Independent Research Ethics Board (CIRB) and the Cojoint Health Sciences Research Ethics Board (CHREB) in Calgary. Participants under the age of majority also completed written informed assent forms and their parent signed a Parent Authorization Form as approved by the research ethics boards. This research complied with the Helsinki Declaration for ethical research. At baseline all consenting participants were assessed using K-SADS/SADS-L format semi-structured interviews conducted by a research psychiatrist (high-risk) or a clinical psychologist (controls). DSM-IV diagnoses were confirmed on blind consensus review using best estimate procedure and taking into account all available research and clinical data as described previously.<sup>9</sup>

Participants were instructed on the study procedure involving weekly and daily remote capture ratings based on True Colours, a web-based platform which, over the last decade, has been used to gather over 1 million self-reports from over 35,000 participants.<sup>1</sup> The True Colours symptom monitoring system uses SMS texting and email to provide an inexpensive and practical means of submitting and reviewing self-reported symptom data. Participants are prompted to complete ratings and a further prompt is sent if a response is not received within 24 hours (for weekly

## High frequency mood ratings in high-risk youth

ratings).<sup>14</sup> All participants were asked to complete weekly ratings and were given the option to complete daily ratings. Participants were asked to complete these ratings for a minimum of 90 consecutive days.

Weekly mood ratings included three validated scales: the General Anxiety Disorder Rating Scale (GADS-7)<sup>15</sup>, the 16-item Quick Inventory of Depressive Symptomatology (QIDS-SR)<sup>16</sup>, and the 5-item Altman Self-Rating Mania Scale (ASRM).<sup>17</sup> Participants were instructed to rate each item according to the description that best captured how they felt over the past 7 days. A weekly email prompt was automatically sent to participants.

Daily Mood Zoom questionnaires were designed to assess day-to-day variations in five positive affect descriptors (i.e., cheerful, excited, optimistic, confident, energetic), five negative affect descriptors (i.e., lonely, anxious, irritated, sad, guilty), and three self-esteem descriptors (i.e., self-liking, shame, self-doubt). Participants were asked “Generally, over today I have felt ...” and rated each descriptor on a scale ranging from 0 (not at all) to 7 (very much). A daily email prompt to enter data was automatically sent to participants.

## Statistical Analysis

Difference in attrition between groups was tested using Cox Proportional Hazard (CPH) models with time to drop-out since study entry as the response and group status as the predictor variable. This method compares the rate of drop-out between the different groups. Compliance was further defined as the proportion of enrolled high-risk offspring and controls completing first 30 and then 90 consecutive days of ratings. The weekly measures were coded as total summary scores and binary variables using the established cut-offs (i.e. 5 for the GADS-7 and 6 for the QIDS-SR and ASRM) as a measure of clinically significant symptoms. Cluster analysis was performed using Principle Components Analysis (PCA) to identify groupings of the daily Mood Zoom questions. The PCA results were then used to obtain a single summary statistic (MZIP1) that accounted for the maximum amount of variability in the 13 daily Mood Zoom responses across all of the subjects in the first 90 days (see Figure S2, Supplementary File). The summary statistic produced from the PCA-MZIP1 was used as a response to test for differences between groups in daily Mood Zoom ratings.

## High frequency mood ratings in high-risk youth

To test for differences between groups over time, linear mixed regression models were used with autoregressive (AR(1)) covariance adjustments for repeated measurements for continuous outcomes, and logistic generalized estimating equations (GEE) were used for binary outcomes. For some of the linear models, the normality assumption was violated and a natural log transformation was performed. In some cases, zero scores were possible and in such cases log (score+1) transformations were performed. In both the linear models and the logistic GEE models, adjustments for sex, age at the start of the study, season, day of the week, and elapsed time in the study were included as other covariates in the models.

Regularity in weekly symptom summary scores and daily ratings for each subject was estimated using the Root Mean Squared Successive Differences (RMSSD) a measure of variability of changes in scores over time, the Teager-Kaiser Energy Operator (TKEO) which combines amplitude and frequency, and the Information Entropy (entropy)<sup>8</sup> which measures regularity of amplitude (see M1, Supplementary File). Each of the regularity metrics measures a slightly different aspect of the stability of the scores over time: high RMSSD signifies large jumps in scores over time, high TKEO signifies more spikes (or dips) in scores over time, and high entropy signifies large overall randomness in scores over time. The within-subject sample variance (Var) was also used as measure of regularity for each subject's responses and estimates the distribution of the amplitudes recorded over time. RMSSD, TKEO, entropy, and Var were all calculated using custom written code in the R programming language version 3.3.2. All statistical analyses were conducted using SAS software version 9.4 for Windows 64bit.

## Results

### *1. Description of the sample*

Fifty high-risk offspring and 108 undergraduate control participants were enrolled in this study. Thirty-one high-risk and 69 controls of the participants agreed to complete daily responses. Fifty-five high-risk offspring were approached for this study and 5 declined to take part owing to lack of interest and/or concern about the required time commitment.

## High frequency mood ratings in high-risk youth

As shown in Table 1 there were more females in both study groups and the mean age of recruitment centered in the early (controls) to mid (high-risk) twenties. The socioeconomic status (SES) of the families, as measured by the Hollingshead, tended to be in the middle to upper middle-class strata for both groups. Recruitment took place across seasons in the high-risk offspring, while in controls recruitment was mostly in the fall-winter season reflecting recruitment over the academic term. In high-risk offspring with a prior mood disorder, 25% met lifetime criteria for a bipolar spectrum disorder (I, II, NOS) and 75% for a depressive spectrum disorder (major depression, depression NOS). All high-risk offspring with a prior mood diagnosis were in good clinical remission at the time of enrolment in this study based on KSADS-PL/SADS-L assessment and symptom rating scales. Specifically, based on log transformed baseline QIDS-SR scores, GADS-7 scores, and ASRM scores, there was no difference between the high-risk offspring without a prior mood disorder compared to those that had a fully remitted prior mood disorder (QIDS-SR  $p=0.3925$ ; GADS-7  $p=0.2561$ ; ASRM  $p=0.2286$ ). Additionally, there was no difference in scores between the two high-risk groups and the control group at baseline (QIDS-SR  $p=0.3315$ ; GADS-7  $p=0.4983$ ; ASRM  $p=0.2075$ ). Furthermore, no high-risk offspring was on maintenance medication. Between high-risk groups with and without a lifetime mood disorder, there were no differences in sex, SES, or season of enrolment; however, those high-risk offspring with a prior mood disorder tended to be older and had a longer duration of weekly and daily ratings.

## *II. Compliance*

With regard to the weekly symptom ratings, 78% (39/50) of high-risk offspring and 97% (105/108) of control offspring completed the first 30 days of data collection, while 56% (28/50) of high-risk and 69% (74/108) of controls completed the full 90 days (see Figure 1). Of the 31 high-risk and 69 control participants who agreed to complete daily Mood Zoom recordings, 77% (24/31) of high-risk and 93% (64/69) of controls completed the first 30 days of data collection while 61% (19/31) of high-risk and 55% (38/69) of controls completed the full 90 consecutive days of data entry. There was no significant difference in the drop-out rate between high-risk and controls within the first 90 days in either of the weekly ( $p=0.2149$ ) or daily ( $p = 0.9792$ ) ratings. Further, there was no difference in rate of dropping out between high-risk offspring with



## High frequency mood ratings in high-risk youth

compared to those without a prior mood diagnosis within the first 90 days of the weekly ( $p=0.1240$ ) or daily ( $p = 0.1329$ ) ratings.

***Insert Figure 1 about here***

### *III. Weekly symptom ratings*

#### *High-risk offspring compared to controls*

Over the first 90 consecutive days, there was no difference in mean weekly GADS-7, QIDS-SR or ASRM summary scores between high-risk offspring and controls, nor was there a difference in the odds of meeting the established cut-offs for each questionnaire (Table S1a; Supplementary File). There was no difference in regularity of weekly ratings over time (RMSSD, TKEO, entropy) between high-risk and controls (Table S1a, Supplementary File). There was evidence that anxiety GADS-7 scores were higher in the fall season ( $p = 0.0010$ ), in females ( $p = 0.0150$ ), and on fridays ( $p = 0.0141$ ). Similarly, season ( $p = 0.0352$ ) and sex ( $p = 0.0288$ ) affected the mean Entropy of GADS-7 scores, and season affected the entropy of QIDS-SR scores ( $p = 0.0075$ ). Specifically, the highest anxiety (GADS-7) entropy was seen in females and the fall season, while depressive (QIDS-SR) entropy was highest in the fall season.

#### *High-risk with and without a prior mood disorder*

There was no difference in average anxiety GADS-7, depressive QIDS-SR, and hypomanic ASRM summary scores in the first 90 days between high-risk offspring with compared to those without a prior mood disorder (Table S1b; Supplementary File). There was marginal evidence ( $p = 0.0590$ ) of a difference in the odds of having hypomanic ASRM scores  $\geq$  the cut-off of 6 in those high-risk with a prior mood disorder. It was estimated that having a prior mood disorder multiplies the odds of Altman scores  $\geq 6$  by 9.63 (95% CI: (0.92, 100.97)).

As shown in Supplementary Table S1b, there was evidence of more irregularity in weekly anxiety GADS-7 scores in high-risk offspring with compared to those without a lifetime mood disorder. Specifically, TKEO-GADS-7 was on average 10.78 units higher ( $p = 0.0030$ ), RMSSD-GADS-7 was 1.1134 higher ( $p = 0.0594$ ) and within-subject variance was 2.843 higher ( $p =$

### High frequency mood ratings in high-risk youth

0.0327). There was marginal evidence of a difference in the regularity of depressive QIDS scores. Specifically, it was estimated that on average, high-risk offspring with a prior mood disorder had higher RMSSD-QIDS ( $p = 0.0628$ ) and higher within-subject variance ( $p = 0.0589$ ). There was no evidence of a difference in the regularity of hypomanic ASRM scores between the high-risk offspring subgroups.

#### IV. Daily Mood Zoom ratings

Cluster analysis based on principle components analysis (PCA) identified 3 clusters, which explained 77.39% of the total variance (Figure S1; Supplementary File). The clusters were as follows: Cluster 1 (positive affect, self-liking) Mood Zoom: q1, q2, q3, q4, q5, q11; Cluster 2 (negative affect): q6, q7, q8, q9; and Cluster 3 (guilt, shame and self-doubt: q10, q12, q13. The first principle component (MZP1) from the PCA was found to account for 56.8% of the total variability. Supplementary Figure S2 illustrates the relative contribution of different questions to the first principal component.

#### *High-risk offspring compared to controls*

There was a significant difference in MZP1 scores between high-risk offspring and controls over 90 days ( $p = 0.0317$ ) as shown in Figure 2. It was estimated that on average, controls had 0.6134 higher MZP1 scores than high-risk offspring, indicating controls had higher daily positive affect/self-liking, lower negative affect, and lower responses in shame and self-doubt. Additionally, mean MZP1 scores were associated with the day of the week ( $p = 0.0452$ ), with Saturday having the highest scores, and Monday having the lowest scores. From Figure 2, it appears that MZP1 adjusted scores may increase over time for high-risk subjects more than for controls, particularly after day 45; however, polynomial and interaction models did not indicate that this difference was significant (all  $p > 0.05$ ).

#### ***Insert Figure 2 about here***

There was no significant difference in the regularity of MZP1 scores as measured by both the mean RMSSD or transformed TKEO between high-risk offspring and controls ( $p = 0.3151$ ;

### High frequency mood ratings in high-risk youth

$p=0.6156$ ). There was some evidence that the season may affect the TKEO of the MZP1 ( $p=0.0412$ ), with fall and winter having more irregular MZP1 scores than spring and summer.

### *High-risk offspring with and without a prior mood disorder*

As illustrated in Figure 3, there was a significant difference in daily MZP1 scores between high-risk offspring with compared to those without a prior mood diagnosis ( $p = 0.0365$ ). It was estimated that on average, high-risk offspring without any prior mood disorder had 1.049 higher MZP1 scores indicating higher positive affect/self-liking, lower negative affect, and lower shame and self-doubt compared to high-risk subjects with a prior but clinically fully remitted mood disorder. This model also indicated mean MZP1 scores were associated with the day of the week ( $p = 0.0346$ ), with Sunday significantly higher than any of the other days of the week.

### *Insert Figure 3 about here*

There was no evidence of a difference in the regularity of MZP1 scores as measured by mean RMSSD ( $p=0.4368$ ) or mean transformed-TKEO MZPC1 ( $p=0.4734$ ) between high-risk offspring with compared to those without any prior mood disorder. There was evidence that season influenced the irregularity of MZP1 as measured by TKEO ( $p = 0.0074$ ), with fall estimated to be the most irregular.

## **Discussion**

In this manuscript we describe a novel prospective 90-day study of self-reported weekly and daily mood-related symptom ratings using a validated remote capture method in the high-risk offspring of bipolar parents compared to university student controls. Main study findings include: (i) evidence of good compliance in both high-risk offspring and control study groups and (ii) the ability to differentiate study groups based on daily ratings of affect and regularity of weekly symptom ratings. Specifically, daily ratings indicated that high-risk offspring experience lower positive affect, higher negative affect, and higher shame and self-doubt compared to controls. High-risk offspring with compared to those without a prior mood disorder had a similar pattern of daily affect ratings and those with a prior mood disorder also had higher irregularity in their weekly ratings of anxiety and depressive symptoms, despite being in full clinical remission.

### High frequency mood ratings in high-risk youth

Taken together these findings support the clinical utility of symptom monitoring using these remote capture self-report methods in this high-risk population.

Over 75% of high-risk offspring and controls completed the data collection using the web-based daily and weekly self-report methods over 30 consecutive days. Complete data rates declined in both groups over the observation period, although at the 90-day mark almost two-thirds of the high-risk offspring had no missing data points in either daily and weekly protocols. These adherence rates are comparable to those reported in similar studies of adult bipolar patients.<sup>8</sup> Feedback from subjects favored the development of an App for smart phones for ease of data entry, particularly when completing recordings on a more frequent (i.e. daily or several times a day) basis. In addition, combining self-reported symptoms with passively collected data using smart phones might also be a fruitful line of inquiry based on findings from a recent study in bipolar adults.<sup>18</sup>

While weekly self-ratings of anxiety, depressive, and hypomanic symptoms did not differentiate the high-risk study group from controls or those high-risk offspring with a prior mood disorder from those without, there was evidence that higher irregularity in mean weekly anxiety and depressive symptom scores did differentiate the high-risk subgroups. Similar challenges in relying on summary scores of self-report mood symptoms to differentiate between study groups or to identify those at risk of mood disorders in high-risk and in the adolescent general population have been described.<sup>6, 19</sup> Therefore, irregularity in self-report symptom ratings in well-defined high-risk individuals may be a more clinical meaningful indicator. Supportive data comes from prior studies in adult bipolar patients which have demonstrated that clinically significant mood symptoms persist during remitted periods<sup>14</sup> and that an increase in irregularity of mood symptoms predicts the recurrence of mood episodes of either polarity.<sup>4</sup>

Over 90 consecutive days, high-risk offspring indicated higher daily levels of negative affect, lower levels of positive affect, and lower self-esteem (shame, self-doubt) compared to controls. Further, there was no difference between high-risk and controls in the regularity of these daily ratings suggesting that these were stable experiences. This observation is consistent with the hypothesis of a psychological vulnerability in young people at familial risk of mood disorder that

## High frequency mood ratings in high-risk youth

might identify a subgroup that would warrant and benefit from prospective symptom monitoring. For example, an independent study reported that high-risk offspring experiencing negative affect responded with rumination, while controls showed both rumination and increased active (adaptive) coping.<sup>20</sup> Lower self-esteem in high-risk offspring but not in controls triggered increased risk taking, while rumination lowered self-esteem at a later time point. In a prior analysis from the Canadian cohort, we found that lower self-esteem predicted a 2.5-fold increased risk of a mood episode and evidence of an interaction effect such that the risk of mood disorder was increased in those offspring with both avoidant coping and lower self-esteem<sup>21</sup>.

In addition to differentiating high-risk offspring from controls, in this study clinically well-remitted high-risk offspring with a prior mood disorder compared to those without reported significantly higher daily levels of negative affect, lower levels of positive affect, and lower self-esteem. Other cross-sectional studies have reported similar findings, namely that offspring of bipolar parents with a prior mood disorder experience lower self-esteem, higher sensitivity to punishment, and rumination in response to negative affect compared to unaffected high-risk offspring.<sup>22, 23</sup> Given that all high-risk subjects in this study were in good clinical remission, findings are consistent with the cognitive vulnerability to depression theory as articulated by Teasdale and colleagues.<sup>24</sup> Specifically, that there may be a persistent psychological vulnerability in high-risk individuals that becomes amplified as a consequence of past mood disorder, and in turn predicts a higher likelihood of future mood episodes. This also aligns with prior findings of the prominence of the depressive polarity of first mood episodes in high-risk offspring in the early stages of emerging bipolar disorder and the possibility that depressive disorders in this population are in fact a part of the bipolar diathesis segregating in relatives of bipolar probands.<sup>25-27</sup>

Despite the promising findings in this study there are limitations that should be discussed. In order to address the question of whether or not frequently reported symptoms will be helpful in predicting mood disorder onset and reducing the time from onset to diagnosis, we would need to observe the subjects longer – ideally through the entire period of risk. However, the aim of this study was to assess compliance with remote capture methods and determine whether or not self-ratings of symptoms using this method might have clinical utility. In this regard, the comparison

### High frequency mood ratings in high-risk youth

between high-risk offspring and controls and between high-risk offspring with compared to those without a prior mood disorder was informative. Secondly, there is a possibility of an attrition bias as controls were given the incentive of completing a mandatory research participation requirement; however, similar completion rates in high-risk and controls suggest differential incentives did not impact our findings. Thirdly, given that the undergraduate controls were self-selected into the study we have no information about those who chose not to participate. However, among the high-risk offspring approached for this study, very few ( $n=5$ ) refused to participate, suggesting acceptable feasibility of these methods in this population. Fourthly, while in good clinical remission based on both clinical assessment and symptom ratings, we cannot prove that the differences in ratings between high-risk subjects with compared to those without a lifetime history of a mood disorder reflect a vulnerability rather than residual symptoms. Strengths of this study include for the first time, the novel application of a remote capture method to a high-risk offspring population. Further, the parental BD diagnosis was confirmed on repeated assessment over many years, ensuring that the high-risk status in the offspring was valid. Further, high-risk offspring have been prospectively studied and repeatedly clinically assessed providing reliable clinical information.

In summary, high-risk subjects showed good rates of compliance with self-report symptom monitoring using remote capture methods. Findings also provide supportive evidence of the clinical utility of symptom monitoring in high-risk youth. Frequent (i.e. daily or several times per day) self-ratings may be informative for identifying symptom patterns that could signal those at highest risk of developing a new or recurrent mood episode. Therefore, high-resolution prospective symptom monitoring using remote capture methods warrant further systematic longitudinal study in well-characterized high-risk populations.

### Funding Sources

This research study was funded by a grant from the Hotchkiss Brain Institute and supported by the National Institute for Health Research and the Oxford Health Biomedical Research Centre. The Flourish high-risk offspring study is funded by a grant from the Canadian Institutes of Health Research PJT 152976.

## High frequency mood ratings in high-risk youth

## References

1. Bopp JM, Miklowitz DJ, Goodwin GM, Stevens W, Rendell JM and Geddes JR. The longitudinal course of bipolar disorder as revealed through weekly text messaging: a feasibility study. *Bipolar Disord.* 2010; 12: 327-34.
2. Bilderbeck AC, Reed ZE, McMahon HC, et al. Associations between mood instability and emotional processing in a large cohort of bipolar patients. *Psychol Med.* 2016: 1-10.
3. Moore PJ, Little, M.A., McSharry, P.E., Goodwin, G.M., Geddes, J.R. Mood dynamics in bipolar disorder. *International journal of bipolar disorders.* 2014; 2: 9.
4. Glenn T, Whybrow PC, Rasgon N, et al. Approximate entropy of self-reported mood prior to episodes in bipolar disorder. *Bipolar Disord.* 2006; 8: 424-9.
5. Hafeman DM, Merranko J, Axelson D, et al. Toward the Definition of a Bipolar Prodrome: Dimensional Predictors of Bipolar Spectrum Disorders in At-Risk Youths. *Am J Psychiatry.* 2016; 173: 695-704.
6. Goodday SM, Preisig M, Gholamrezaee M, Grof P, Angst J and Duffy A. The association between self-reported and clinically determined hypomanic symptoms and the onset of major mood disorders. *BJPsych Open.* 2017; 3: 71-7.
7. Fredrickson B. Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion.* 2000; 14: 577-606.
8. Tsanas A, Saunders KE, Bilderbeck AC, et al. Daily longitudinal self-monitoring of mood variability in bipolar disorder and borderline personality disorder. *J Affect Disord.* 2016; 205: 225-33.
9. Duffy A, Horrocks J, Doucette S, Keown-Stoneman C, McCloskey S and Grof P. The developmental trajectory of bipolar disorder. *The British journal of psychiatry : the journal of mental science.* 2014; 204: 122-8.
10. Turecki G, Grof P, Cavazzoni P, et al. Evidence for a role of phospholipase C-gamma1 in the pathogenesis of bipolar disorder. *Mol Psychiatry.* 1998; 3: 534-8.
11. Grof P, Alda M, Grof E, Zvolsky P and Walsh M. Lithium response and genetics of affective disorders. *J Affect Disord.* 1994; 32: 85-95.
12. Grof P, Duffy A, Cavazzoni P, et al. Is response to prophylactic lithium a familial trait? *The Journal of clinical psychiatry.* 2002; 63: 942-7.
13. Grof P, Duffy A, Alda M and Hajek T. Lithium response across generations. *Acta Psychiatr Scand.* 2009; 120: 378-85.
14. McKnight RF, Bilderbeck AC, Miklowitz DJ, Hinds C, Goodwin GM and Geddes JR. Longitudinal mood monitoring in bipolar disorder: Course of illness as revealed through a short messaging service. *J Affect Disord.* 2017; 223: 139-45.
15. Spitzer RL, Kroenke K, Williams JB and Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006; 166: 1092-7.
16. Trivedi MH, Rush AJ, Ibrahim HM, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med.* 2004; 34: 73-82.
17. Altman EG, Hedeker D, Peterson JL and Davis JM. The Altman Self-Rating Mania Scale. *Biol Psychiatry.* 1997; 42: 948-55.

## High frequency mood ratings in high-risk youth

18. Faurholt-Jepsen M, Vinberg M, Frost M, Christensen EM, Bardram J and Kessing LV. Daily electronic monitoring of subjective and objective measures of illness activity in bipolar disorder using smartphones--the MONARCA II trial protocol: a randomized controlled single-blind parallel-group trial. *BMC Psychiatry*. 2014; 14: 309.
19. Tijssen MJ, van Os J, Wittchen HU, et al. Prediction of transition from common adolescent bipolar experiences to bipolar disorder: 10-year study. *Br J Psychiatry*. 2010; 196: 102-8.
20. Pavlickova H, Turnbull OH, Myin-Germeys I and Bentall RP. The inter-relationship between mood, self-esteem and response styles in adolescent offspring of bipolar parents: an experience sampling study. *Psychiatry research*. 2015; 225: 563-70.
21. Goodday S, Bentall, R., Jones, S., Weir, A., Duffy, A. Coping strategies and self-esteem in the high-risk offspring of bipolar parents. *Australian and New Zealand Journal of Psychiatry*. 2018; In Press.
22. Pavlickova H, Turnbull O and Bentall RP. Cognitive vulnerability to bipolar disorder in offspring of parents with bipolar disorder. *The British journal of clinical psychology / the British Psychological Society*. 2014.
23. Jones SH, Tai S, Evershed K, Knowles R and Bentall R. Early detection of bipolar disorder: a pilot familial high-risk study of parents with bipolar disorder and their adolescent children. *Bipolar Disord*. 2006; 8: 362-72.
24. Teasdale JD, Segal Z and Williams JM. How does cognitive therapy prevent depressive relapse and why should attentional control (mindfulness) training help? *Behav Res Ther*. 1995; 33: 25-39.
25. Duffy A, Alda M, Hajek T and Grof P. Early course of bipolar disorder in high-risk offspring: prospective study. *The British journal of psychiatry : the journal of mental science*. 2009; 195: 457-8.
26. Duffy A. Early identification of recurrent mood disorders in youth: the importance of a developmental approach. *Evidence-based mental health*. 2015; 18: 7-9.
27. Blacker D, Faraone SV, Rosen AE, et al. Unipolar relatives in bipolar pedigrees: a search for elusive indicators of underlying bipolarity. *American journal of medical genetics*. 1996; 67: 445-54.



## High frequency mood ratings in high-risk youth

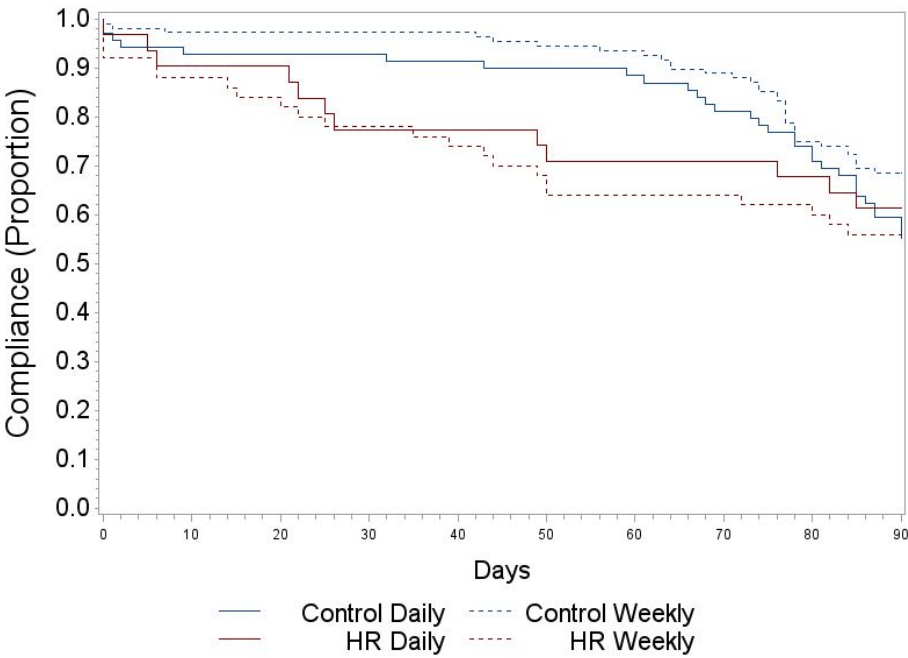
Table 1 Description of study groups

	High-risk offspring (n = 50)	Control offspring (n = 108)	p-value	High-risk without prior mood disorder (n = 28)	High-risk with prior mood disorder (n = 22)	p- value
<b>Sex n(%)</b>						
Male	12 (24.0)	13 (12.4)	0.0660 <sup>a</sup>	9 (32.1)	3 (13.6)	0.1283 <sup>a</sup>
Female	38 (76.0)	92 (87.6)		19 (67.9)	19 (86.4)	
<b>SES n(%)</b>						
1	0 (0)	0 (0)	0.0036 <sup>b</sup>	0 (0)	0 (0)	0.1085 <sup>b</sup>
2	0 (0)	5 (6.0)		0 (0)	0 (0)	
3	1 (2.0)	12 (14.5)		0 (0)	1 (4.5)	
4	27 (55.1)	25 (30.1)		18 (66.7)	9 (40.9)	
5	21 (42.9)	41 (49.4)		9 (33.3)	12 (54.6)	
<b>Season at Enrollment n(%)</b>						
Spring	14 (28.0)	0 (0)	<0.0001 <sup>b</sup>	10 (35.7)	4 (18.2)	0.5973 <sup>b</sup>
Summer	10 (20.0)	1 (0.9)		5 (17.9)	5 (22.7)	
Fall	7 (14.0)	63 (58.3)		3 (10.7)	4 (18.2)	
Winter	19 (38.0)	44 (40.7)		10 (35.7)	9 (40.9)	
<b>Age at Enrolment</b>						
Mean (SD)	26.20 (6.40)	21.01 (2.24)	<0.0001 <sup>c</sup>	23.78 (6.04)	29.28 (5.55)	0.0018 <sup>c</sup>
<b>Daily Response Duration in Days</b>						
Mean (SD)	225.60 (207.6)	173.00 (150.8)	0.2114 <sup>c</sup>	168.40 (203.6)	266.90 (206.2)	0.1971 <sup>c</sup>
<b>Weekly Response Duration in Weeks</b>						
Mean (SD)	27.33 (2837)	26.17 (20.66)	0.7957 <sup>c</sup>	21.07 (24.94)	35.29 (30.97)	0.0783 <sup>c</sup>

<sup>a</sup>Chi-square test<sup>b</sup>Fisher's exact test<sup>c</sup>t-test

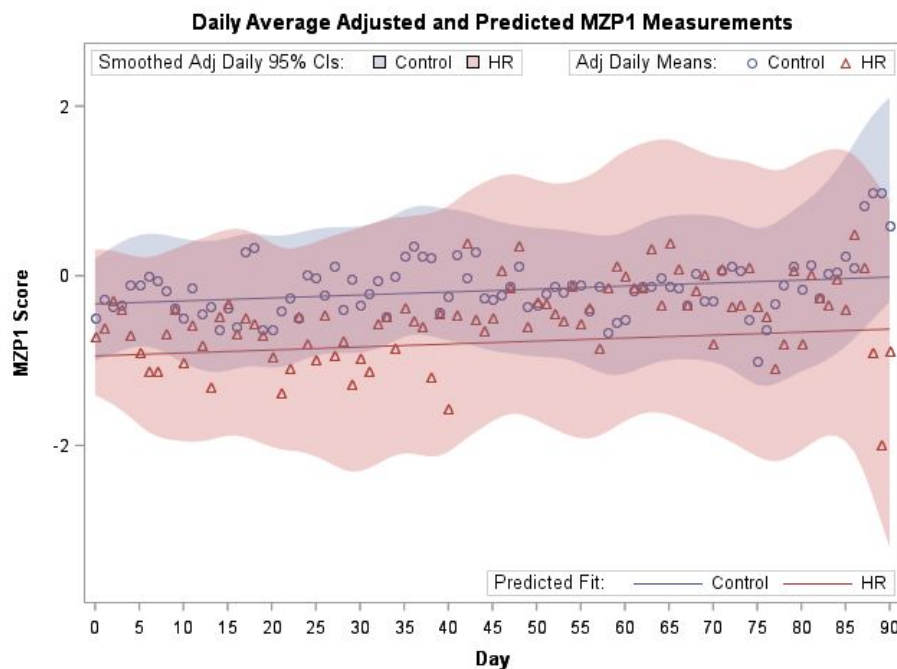
High frequency mood ratings in high-risk youth

Figure 1. 90-Day compliance in high-risk (HR) offspring and controls for weekly and daily ratings



## High frequency mood ratings in high-risk youth

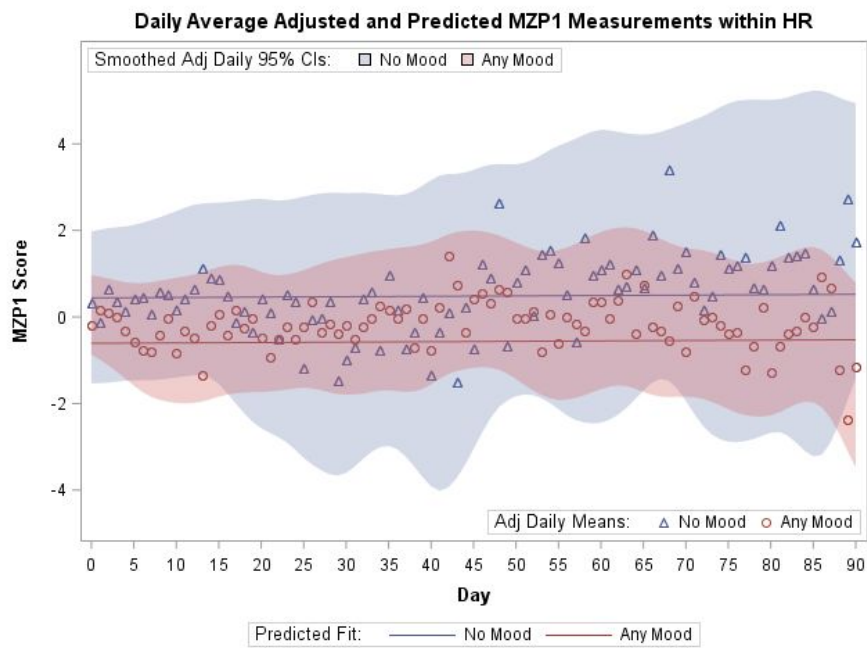
Figure 2. Daily average MZP1 affect symptom ratings in high-risk offspring vs control



Legend: the scatter plot shows the control and high-risk daily means of the MZP1 scores adjusted for variability due to day of the week, sex, season, and age. The translucent bands show the smoothed 95% confidence intervals for each of the daily means of the adjusted scores. The solid straight lines show the linear mixed model predicted mean value of adjusted MZP1 score for the two groups

High frequency mood ratings in high-risk youth

Figure 3. Daily average MZP1 affect symptom ratings in high-risk offspring with compared to those without prior mood disorder



Legend: the scatter plot shows the daily means of the MZP1 scores for high-risk with any diagnosable mood compared to those without after accounting for variability due to day of the week, sex, season, and age. The translucent bands show the smoothed 95% confidence intervals for each of the daily means of the adjusted scores. The solid straight lines show the linear mixed model predicted mean value of adjusted MZP1 score for the two groups.

## Supplementary File

Table S1a P-values comparing True Colours weekly mood ratings over 90 days in high-risk offspring and controls

	GADS-7	QIDS-SR	ASRM
<b>Mean Score</b> ( $\beta_{HR - C}$ )	0.4027† (0.104)	0.2645† (0.126)	-
<b>Odds (cut-off)</b> ( $OR_{HR/C}$ )	0.9557 ( $\geq 5$ ) (0.979)	0.6445 ( $\geq 6$ ) (1.207)	0.8002 ( $\geq 6$ ) (0.870)
<b>RMSSD</b> ( $\beta_{HR - C}$ )	0.3454 (0.353)	0.7912 (0.098)	0.3737 (-0.302)
<b>TKEO</b> ( $\beta_{HR - C}$ )	0.6666 (0.904)	0.8037 (-0.524)	0.5631† (-0.100)
<b>Entropy</b> ( $\beta_{HR - C}$ )	0.3575 (0.110)	0.6337 (-0.053)	0.7929 (-0.0392)
<b>Log Var</b> ( $\beta_{HR - C}$ )	0.6792 (0.131)	0.4909 (0.205)	0.3358 (-0.351)

†Transformation of response required

\*p<0.05

Table S1b P-values comparing True Colour Weekly Mood Ratings in high-risk offspring with prior mood disorder compared to those without over 90 days

	GADS-7	QIDS-SR	ASRM
<b>Mean Score</b> ( $\beta_{With\ Mood - Without}$ )	0.1114 (0.296)	0.1571† (0.254)	-
<b>Odds (cut-off)</b> ( $OR_{With\ Mood/Without}$ )	0.2378 ( $\geq 5$ ) (2.074)	0.5043 ( $\geq 6$ ) (1.586)	0.0590 ( $\geq 6$ ) (9.625)
<b>RMSSD</b> ( $(\beta_{With\ Mood - Without})$ )	0.0594 (1.113)	0.0628 (0.791)	0.2799 (0.379)
<b>TKEO</b> ( $(\beta_{With\ Mood - Without})$ )	0.0030* (10.776)	0.1556 (2.697)	0.2881† (0.298)
<b>Entropy</b> ( $(\beta_{With\ Mood - Without})$ )	0.0902 (0.253)	0.1122 (0.249)	0.1526 (0.284)
<b>Log Var</b> ( $(\beta_{With\ Mood - Without})$ )	0.0327* (1.045)	0.0589 (0.764)	0.7177 (0.211)

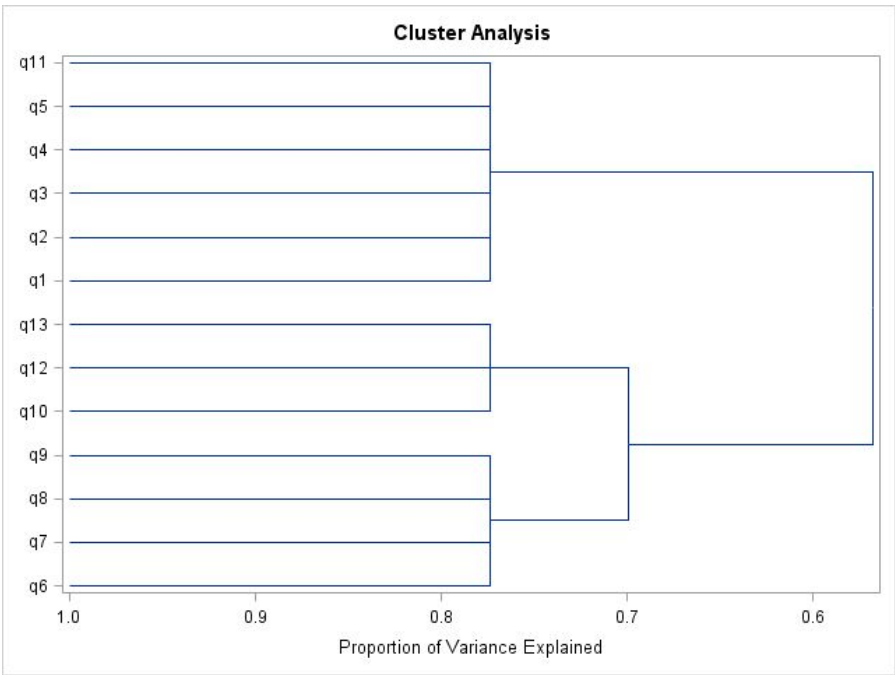
†Transformation of response required

\*p<0.05

Supplementary File

**Results from the cluster and principle component analysis of daily affect ratings**

Figure S1. Proportion of variance explained by each cluster



From the cluster analysis, the identified clusters were as follows:

Postive Affective Cluster 1: q1, q2, q3, q4, q5, q11

Negative Affective Cluster 2: q6, q7, q8, q9

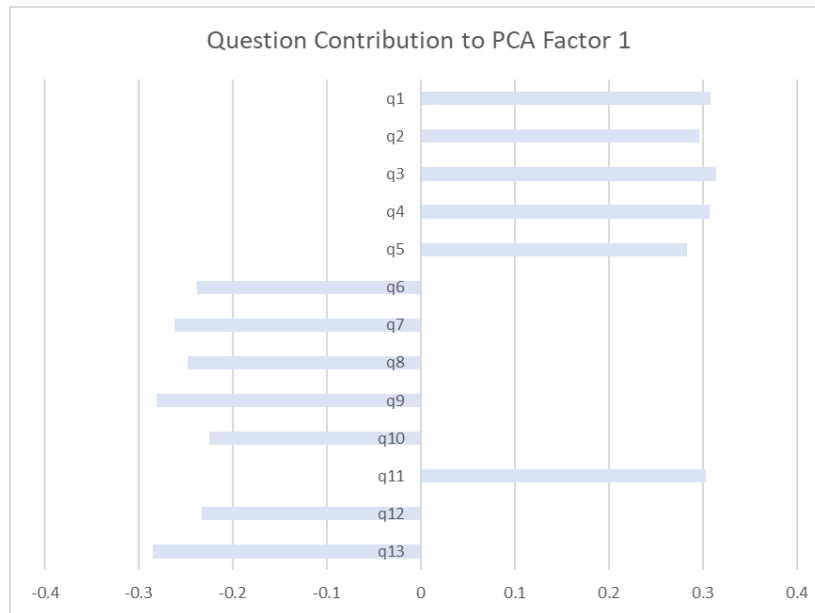
Self-Esteem Cluster 3: q10, q12, q13

## Supplementary File

**Construction of a single summary score using principle component analysis**

The first principal component (MZP1) was found to account for 56.78% of the total variability (the second, unused component accounted for an additional 13.80%). The plot below shows the relative contribution of the different questions to the first principal component (after standardization of each question).

Figure S2. Contribution of each daily mood zoom question to the single summary score (MZP1)



## Supplementary File

M1 Statistical Model Regularity

Three different measures of regularity of scores were performed: 1) the root mean squared successive differences (RMSSD); 2) the Teager-Kaiser Energy Operator (TKEO); and 3) information entropy (entropy).

For  $n$  measurements of  $x_1, x_2, \dots, x_n$  over time, RMSSD, TKEO, and entropy were calculated as follows,

$$RMSSD_X = \sqrt{\frac{1}{n} \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2}$$

$$TKEO_X = \frac{1}{n} \sum_{i=2}^{n-1} (x_i^2 - x_{i-1} \cdot x_{i+1})$$

$$Entropy_X = \sum_{i=1}^n [P(x_i) \cdot \log (P(x_i))]$$

Note that the given form of information entropy cannot be applied to continuous responses.