Audit and feedback of therapist-assisted internet-delivered cognitive behaviour therapy within routine care: A quality improvement case study

Heather D. Hadjistavropoulos\textsuperscript{a,⁎}, Jaime Williams\textsuperscript{b}, Kelly Adlam\textsuperscript{c}, Kerry Spice\textsuperscript{d}, Marcie Nugent\textsuperscript{a}, Katherine M.B. Owens\textsuperscript{b}, Christopher Sundström\textsuperscript{b}, Blake F. Dear\textsuperscript{c}, Nickolai Titov\textsuperscript{d}

\begin{abstract}
With the growing use of ICBT in routine care clinics there is a need for literature on how to monitor and improve the quality of therapist behaviours in clinical practice. In this paper, we first provide background literature on Audit and Feedback (A&F), a common quality improvement technique, and then present a case study regarding the use of A&F to improve quality of therapist behaviours in emails sent to patients provided with ICBT in routine care. The A&F measure used was derived from previous research on therapist's email behaviours in ICBT. Fifteen undesirable therapist behaviours (e.g., Did Not Message, Unresponsive to Symptom Increase, Does Not Address Patient Concern) were audited in 1840 emails sent from eight therapists to 198 randomly selected patients, representing 18% of 1114 patients who started between one and five lessons of ICBT in the previous year and did not formally withdraw from treatment (n = 31 patients). The therapists who were audited were provided feedback four times over a one-year period from October 2018 to September 2019. Overall, in all audit periods, we found a low percentage of undesirable therapist behaviours (i.e., therapists displayed the behaviour in 12% or less of the total emails sent). For most therapist behaviours, we saw a trend towards improvement across the four audit cycles. Three therapist behaviours (i.e., Failure to Ask One Question to the Patient, Poor Instructions, Not Linking Email to Course Content) did not follow this pattern and were flagged for clinical discussion to determine why behaviours were elevated and whether these behaviours represented unrealistic expectations. The process was valuable for monitoring and improving therapist behaviours and highlights the need for future research on standards for therapist behaviours (e.g., which behaviours to focus on, setting acceptable levels of undesirable behaviour).
\end{abstract}

1. Introduction

Internet-delivered Cognitive Behaviour Therapy (ICBT) is effective for ameliorating symptoms of many different mental health disorders, including depression and anxiety (e.g., Andersson et al., 2019; Carlbring et al., 2018). Given the strength of this research evidence, as well as the practical benefits of ICBT (e.g., increased access to care regardless of location, time and mobility), there has been growing attention to implementation of ICBT in routine care (Titov et al., 2018).

When ICBT is offered in routine care, some form of therapist contact is usually provided, either via real-time (e.g., telephone or chat), delayed interaction (e.g., secure e-mail), or a combination of these (e.g., Andersson, 2016). As ICBT has evolved from existing primarily in research settings to now also being used in routine care (Titov et al., 2018), there is a need for methods to monitor and improve the quality of therapist behaviours. Such methods are needed, not only to safeguard patients, but also to ensure retention of the benefits identified in research trials. As noted by Perepletchikova (2011), “once an intervention is validated, its utilization with high integrity may increase chances for outcomes similar to those in the original efficacy research” (p. 148). To date, few studies have investigated quality improvement methods for monitoring and improving quality of therapist behaviours.

\begin{keywords}
Internet-delivered cognitive behaviour therapy \\
Audit and feedback \\
Quality improvement \\
Therapist behaviours
\end{keywords}

\begin{abbreviations}
A&F, Audit & Feedback; AC, Audit cycle; ICBT, Internet-delivered cognitive behaviour therapy; ICBT-UTBS, ICBT Undesirable Therapist Behaviours Scale; ICBT-TRS, ICBT Therapist Rating Scale
\end{abbreviations}

⁎ Corresponding author.

\begin{e-mail}
hadjista@uregina.ca (H.D. Hadjistavropoulos), Jaime.williams@uregina.ca (J. Williams), Kelly.adlam@uregina.ca (K. Adlam), Kerry.spice@uregina.ca (K. Spice), Marcie.Nugent@uregina.ca (M. Nugent), katherine.Owens@saskhealthauthority.ca (K.M.B. Owens), Christopher.Sundstrom@uregina.ca (C. Sundström), blake.dear@mq.edu.au (B.F. Dear), nick.titov@mq.edu.au (N. Titov).
\end{e-mail}

https://doi.org/10.1016/j.invent.2020.100309

Received 10 December 2019; Received in revised form 17 January 2020; Accepted 25 January 2020

© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
in ICBT. In this paper, we first provide an overview of audit and feedback (A&F), a well-established and convenient technique for quality improvement used in healthcare settings (e.g., Ivers et al., 2012; Pedersen et al., 2018a), and then outline our application of A&F within a routine care ICBT clinic.

1.1. Audit and feedback

The overall purpose of A&F is to measure care provided by an organization against relevant and recognized standards and to correct identified problems or areas that are deficient or lacking (Aveline and Watson, 2001). A&F involves a feedback loop, wherein an audit is first conducted, after which commentary on performance is provided to practitioners who then target the problematic issues. This is followed by further identification of problems (i.e., re-audit), feedback, and change (Pedersen et al., 2018a). Relevant standards are derived through clinical experience or previous research and are assessed and analyzed in a systematic way (National Institute for Clinical Excellence, 2002). The aim of A&F is to promote new positive changes in the measured outcomes with each subsequent audit, so that there is a continuous flow of improvement in both team performance and patient outcomes (Aveline and Watson, 2001; McPherson and Richardson, 2018).

Although A&F has been heralded as one of the most effective behavioural interventions for improving care (Hysong et al., 2012), some descriptions are more cautious. In a Cochrane review, Ivers et al. (2012) reported that although A&F may lead to important improvements in clinical practice, there is significant variability in outcomes among studies. Improvements appear dependent, in part, upon the means by which feedback is provided to professionals. For instance, feedback has been found to be more valuable when delivered by a trusted source (such as a colleague), when delivered in two or more modalities (such as both discussion and text), and when comparison with peers is included (Ivers et al., 2014). It is also helpful if the feedback leads to specific goals and plans. Pedersen et al. (2018a) note that in mental health care, ambivalence, lack of support, and competing demands that are given greater weight can contribute to lack of improvement with A&F. Further, the outcome of A&F also relates to the level of baseline clinical performance; those who have a large amount of improvement to make, tend to show greater progress compared to those who begin at a level closer to their target (Ivers et al., 2012).

Colquhoun et al. (2017b) conducted a review of A&F interventions to investigate why A&F is effective in some contexts but not in others. Using consensus-based methodology, the authors identified 17 elements that impact effectiveness of A&F. These elements were grouped into the following six categories: 1) whom (to whom the audit was targeted); 2) what (what audited information was provided as feedback); 3) why (why this information was given); 4) when (when the feedback was provided); 5) how (how the feedback was given) and; 6) how much (how much feedback was given). The first category, Who, had two elements associated with it: whether the feedback was intended for individuals or groups and whether it was given in person to whom the behavioural change was required. The second category, What, comprised elements of the process of care, patient outcomes, individual provider performance, individual patient cases, specific behaviours needed to change, graphical elements presented with the feedback, and comparison of performance with others. The third category, Why, consisted of one element related to a rationale being provided for the audit. The fifth category, How, had two items related to whether the feedback was provided face-to-face and if providers were asked to consider whether the changes would impact their practice. The sixth category, How Much, contained the element of number of times feedback was provided. Colquhoun et al. (2017b) recommended that those considering implementing A&F for quality improvement purposes should attend to all 17 elements.

One recent example of how A&F may be used in mental healthcare is provided by Pedersen et al. (2018b) who conducted an audit of staff from different disciplines including social work, nursing and psychology. They describe undertaking a full audit with a re-audit occurring approximately one year later in order to examine whether staff were offering care congruent with Norwegian national guidelines on mental health and addiction. Practice was audited with feedback given both verbally and in writing about four weeks after the audit. The audit was used by the staff to develop action plans to improve various aspects of care, such as screening of alcohol and drug use and treatment of alcohol and drug problems concurrent with mental illness. Overall, Pedersen et al. (2018b) found that although the audit process was determined by upper management, the clinical staff were enthusiastic about the process. Nevertheless, staff felt that there were significant barriers to implementation noting that while many action plans may have been initially drawn up with eagerness, these would often dissolve with little action. Factors such as competing workplace responsibilities and not having a designated leader contributed to this. Ultimately, the participants did not report feeling ownership over the A&F cycle, questioning management’s actual interest in A&F.

1.2. Case study purpose

The purpose of the current case study is to provide results of an A&F of therapists delivering ICBT in routine care. Following the framework provided by Colquhoun et al. (2017b), our A&F focused on therapists who delivered ICBT (whom) and specifically attended to breaks in practice or undesirable therapist behaviours informed by past literature on therapist behaviours in ICBT (what). Cases were audited regularly and feedback was provided to the therapists each week when warranted and as an overall index every three months (when). Feedback occurred in multiple ways (how, i.e., individual feedback, group email, group supervision) for a total of four cycles (how much). In conducting the A&F project, we were interested in answering the following questions:

1) To what extent would undesirable therapist behaviours be present in emails among trained therapists in clinical practice?
2) Would undesirable therapist behaviours in emails improve over the course of A&F cycles and, if so, which ones?

Based on past research (Hadjistavropoulos et al., 2019a), we expected some variability in undesirable therapist behaviours, but we also expected that the majority of emails would not display undesirable therapist behaviours. Based on the A&F literature, we hypothesized that we would see improvement in undesirable therapist behaviours through the A&F process.

2. Method

2.1. Therapists

Therapists involved with the audit delivered free ICBT via the Online Therapy Unit, a routine ICBT clinic funded by the province of Saskatchewan. The clinic is primarily focused on delivering a trans-diagnostic ICBT course for anxiety and depression (the Wellbeing Course described below), but also offers ICBT for alcohol misuse, chronic pain and chronic health conditions. This paper focusses specifically on the audit of therapists providing the Wellbeing Course. Patients whose cases were audited were assigned to one of four therapists employed by the Online Therapy Unit, or to one of four therapists trained in ICBT and employed in a publicly-funded community mental health clinic. Previous research comparing provision of the Wellbeing Course by therapists in the ICBT clinic with therapists at the community mental health clinic has not found differences in completion rates, satisfaction or clinical outcomes of patients (Hadjistavropoulos et al., 2016a).

The total number of therapists audited for this project was eight, although the number varied by audit cycle (AC). In AC1 (October 1–December 31, 2018), six therapists provided care and thus, their cases were audited. This number rose to seven in AC2 (January 1, 2019–March
31, 2019), and then to eight during AC3 (April 1, 2019–June 30, 2019), a number that was maintained during AC4 (July 1, 2019–September 30, 2019). In total, during the one-year audit period, 1114 patients were eligible for audit in that they were provided with the Wellbeing Course and started between 1 and 5 lessons. There were an additional 31 (2.7%) patients who formally withdrew from treatment and research and were therefore not eligible for audit; therapists had a similar number of patients who formally withdrew from treatment (0%, 0.65%, 1.35%, 2.69%, 3.13%, 4.07%, 4.41%, 4.74%). In each cycle, the auditor was requested to randomly review ~20% of cases that were eligible for audit and had completed involvement in the unit during the AC. This percentage was chosen because it was deemed large enough for problems to be detected yet small enough for the workload to be managed. As therapists had variable caseloads, and cases were selected randomly, the number of cases reviewed per therapist ranged from three to 11 during AC1 (n = 39); three to 10 during AC2 (n = 40); five to 15 during AC3 (n = 76); and three to seven during AC4 (n = 43). In the end, the percentage of cases audited was 18% (198/1114) rather than 20% of cases as some weeks no audits were conducted due to holidays. As patients varied in treatment engagement (e.g., patients started between 1 and 5 lessons) and as therapists also do not send messages when they were on holidays, the number of messages audited from therapist to patient ranged from six to 20 (M = 9.67, SD = 3.31; total 377 messages) for AC1, from six to 11 (M = 8.55, SD = 1.13; total 342 messages) for AC2, from six to 17 (M = 8.96, SD = 2.06; total 681 messages) for AC3, and from one to 18 (M = 10.23, SD = 4.17; total 440 messages) for AC4.

2.2. Screening and intervention

All patients entered ICBT through a common screening process. Patients first completed online measures, querying demographic information, basic inclusion criteria (e.g., access to computer and internet, emergency contact), symptom severity (i.e., the Patient Health Questionnaire-9; PHQ-9; Kroenke et al., 2001, and the Generalized Anxiety Disorder-7; GAD-7; Spitzer et al., 2006), alcohol and drug use, as well as past mental health diagnoses and treatment. To be eligible for the course, patients had to be 1) 18 years of age or older, 2) a Saskatchewan resident, 3) endorsing at least mild symptoms of depression and/or anxiety (PHQ-9 and GAD-7 > 4), 4) willing to provide an emergency medical contact, and 5) have access to a computer and the internet. Following this online screening, telephone contact was made with the prospective patient for further screening. Referral was made to a face-to-face practitioner if the person was at high risk of suicide, had significant concerns with alcohol or drugs, or a severe mental health diagnosis that was unmanaged (e.g., psychosis, bipolar). Patients who were deemed appropriate for ICBT were provided with a username and password giving them access to the Wellbeing Course.

The Wellbeing Course was developed by the eCentreClinic (www.ecentreclinic.org) at Macquarie University, Sydney, Australia (Titov et al., 2011, 2015), and previous research has demonstrated its efficacy in decreasing symptoms of depression, anxiety, distress, and disability (e.g., Hadjistavropoulos et al., 2017; Hadjistavropoulos et al., 2016b; Hadjistavropoulos et al., 2019b; Titov et al., 2011, 2015). Details regarding the context of the five lessons of the Wellbeing Course are available found by reviewing Titov et al. (2015). In brief, patients work through the following five online lessons with suggested homework, which are gradually released over an eight-week period and address: 1) identifying symptoms and the cognitive behavioural model; 2) monitoring and challenging thoughts; 3) de-arousal strategies and scheduling pleasant activities; 4) graduated exposure; and 5) relapse prevention. Patients complete measures of depression and anxiety prior to each of the five core lessons. Patients are also provided with reminder email messages informing them about the availability and content of upcoming lessons.

2.3. Therapist assistance

Each patient is assigned to a designated therapist. On a weekly basis, the assigned therapist logs in to review patient responses to questionnaires, track patient progress on materials, and securely review messages from and send messages to patients. Telephone contact is provided if the patient does not log on during the previous week, requests such contact, or has a substantial increase in symptoms (≥5 points on the PHQ-9 or GAD-7).

All therapists receive a one-day workshop (Hadjistavropoulos et al., 2012) addressing research and professional practice related to ICBT. They also complete self-directed lessons on ethics, the history and research findings of the Online Therapy Unit, as well as the Wellbeing Course. This training is followed by supervised practice on all cases, for at least 10 weeks until therapists are identified as consistently exhibiting the following therapist behaviours in their emails: Builds Rapport, Seeks Feedback, Provides Symptom Feedback, Provides Psychoeducation, Facilitates Understanding, Praises Effort, Encourages Practice, Clarifies Administrative Procedures, and Communicates Effectively (Hadjistavropoulos et al., 2018b). The official treatment period is 8 weeks although some final exchanges may occur after this period in order to provide patients with final feedback on symptom improvement. All contact with patients is expected to be completed by 10 weeks.

2.4. Audit

The audit form (see Appendix 1) was developed by several members of the Online Therapy Unit (HDH, KA, KS, MN) and piloted and refined (e.g., definitions were refined for each behaviour to ensure consistency in use of the form) by two senior therapists (KA, KS) over a six-month period prior to the A&F process (see Appendix 1). Specifically, the two auditors scored the same set of three randomly selected patients. The auditors checked their ratings with each other, discrepancies were resolved verbally and the definitions for the items were refined. This process was repeated three times, after which auditors were obtaining the same scores. However, no formal type of inter-rater reliability was calculated. The two auditors who piloted and refined the audit form, continued to conduct the audit described in this paper (KA completed 123 audits and KS 75).

The audit form was developed to efficiently and systematically identify areas of emails that were lacking, that could result in poor patient engagement and outcomes, and that could be targeted for improvement. The audit form has much in common with the ICBT Undesirable Therapist Behaviours Scale (ICBT-UTBS; Hadjistavropoulos et al., 2019a) which was developed by the Online Therapy Unit to identify the frequency, nature and correlates of undesirable therapist behaviours. The ICBT-UTBS assesses six undesirable therapist behaviours that were initially identified through qualitative analysis of emails (Hadjistavropoulos et al., 2019a). Undesirable therapist behaviours include: inadequate detail in emails to the patient; failure to address patient content; not being supportive; not providing correspondence when warranted; disclosing personal information inappropriately and; failure to appropriately manage risk.

All items from the ICBT-UTBS are captured on the audit form, but have been refined in order to be able to provide therapists with constructive feedback (see Appendix 1). For example, the ICBT-UTBS therapist behaviour “failure to address content” was broken down into Does Not Answer Patient Question and Does Not Address Patient Concern. The therapist behaviour “Inadequate detail” was broken down into Lack of Psychoeducation, Failing to Ask Patients One Question, Poor Instructions and Not Linking Patient Concerns to Course Content. In addition to assessing “Did not message”, the audit form was also expanded to capture other basic practices that are important in the unit, such as Did Not Call as Indicated, No Contact Note and Did Not Provide Patient with Next Check-in Date. Finally, the audit form also captured Poor Writing and Poorly Timed Emails. The form was subsequently divided into four basic Practice items such as No Contact Note and Did Not Message and 11 Content Audit items such as Lack of Psychoeducation and Did Not Link to Course Content.

Each audit assessed patients’ full treatment period, with number of emails varying to some extent depending on patient engagement, whether therapists had holidays and whether patients had a clinical
issue (e.g., crisis) requiring more than once a week emails (M = 9.29; SD = 2.87, range 1–20 emails). Each therapist behaviour was scored as either “present” or “absent” for each email. At the end of each audit form, auditors were prompted to identify items for supervision (coaching topics). In A&F, it is common to set targets based on clinical judgment (see Colquhoun et al., 2017b). As a matter of clinical practice, we chose 5% or greater to indicate a level of undesirable therapist behaviours that warranted attention.

2.5. Feedback

One auditing therapist (KA) was responsible for providing feedback and coaching to all audited therapists. Over time, an increasing amount of feedback was provided to therapists. Throughout the entire period, individual therapists were contacted by phone and/or email about any therapist behaviours that were identified as being of significant concern. For example, when scores were significantly elevated (i.e., 4 or more emails to the patient were identified as having the same undesirable behaviour or there were a total 10 or more instances of problematic behaviours), the primary auditing therapist provided verbal feedback to therapists suggesting they work on specific areas. In these cases, therapists were sent samples of quality work and clarification of therapist behaviours was provided.

In terms of group feedback, after all ACs, the auditing therapist provided general feedback to all therapists in a group supervision meeting about the audit findings. Specifically, group meetings focussed on therapist behaviours that as a group did or did not improve (e.g., the AC1 meeting focussed on the need for therapists as a group to attend to symptom scores, address patient concerns, and improve instructions to patients; the AC2 meeting addressed these issues and also discussed the importance of therapist’s writing clear emails and directing patients to specific course content; the AC3 meeting focussed on the need to attend to scores on measures and quality of instructions; the AC4 meeting focussed on discussion of failing to ask at least one question to patients and poor instructions). Of note, scores were not shared at the group meetings. Instead, general feedback was provided to therapists. Throughout the entire period, individual therapists were sent individual reports that included information on each patient that demonstrated the undesirable therapist behaviour), and the percentage of all emails across all patients that were audited that demonstrated the undesirable therapist behaviour. The supervisor of therapists working at the publicly-funded community mental health clinic also received this information (KMBO) and was requested to follow up with therapists on any higher frequency undesirable therapist behaviours (e.g., discuss and problem solve how to improve undesirable therapist behaviours that were identified in 5% or more of the emails). Similarly, the auditing therapist also did this for therapists working out of the Online Therapy Unit. For the third and fourth AC, each therapist was asked to set individual goals for the next AC (e.g., if a score was above 5%, the therapist was asked to work on that behaviour for the next AC).

3. Results

3.1. Per audit cycle

Results per AC are presented in Table 1. For all ACs (1–4), mean percentages of undesirable therapist behaviours were 12% or less. Nevertheless, as apparent from examination of the SDs, there was significant variability among therapists in presence of undesirable therapist behaviours in emails. The highest percentage scores were all within AC1. During AC1, the highest percentage score was for Unresponsive to symptom increase, scoring measures, or suicide risk. This typically reflected not commenting on a symptom change rather than not responding to suicide risk. During AC2, the highest percentage score was for the same item; again upon review, this reflected therapists not commenting to a symptom change rather than suicide risk. During AC3 the highest percentage score was for the behaviour, Poor Instructions. During AC4, the highest percentage score was for the behaviour Failing to Ask One Question. Concerning the lowest percentages, these were for Unnecessary Self-Disclosure, No Contact Note, and Does not Answer Patient Question.

For many of the items, there were trends towards improvement across ACs or an indication that scores were low throughout all ACs (see Table 1). Six out of 15 behaviours were never above 5% during an audit period; another 5 of the 15 behaviours were above 5% but declined

| Audit cycle 1 | Audit cycle 2 | Audit cycle 3 | Audit cycle 4 |
|---------------|---------------|---------------|---------------|
| M% (SD) | M% (SD) | M% (SD) | M% (SD) |
| N cases = 39 | N cases = 40 | N cases = 76 | N cases = 43 |
| Basic practice | | | |
| Did not message | 6.91 (10.60) | 5.41 (10.88) | 3.51 (9.39) | 1.42 (4.62) |
| Did not call as indicated | 5.33 (10.79) | 4.63 (11.95) | 1.07 (3.42) | 1.34 (5.69) |
| No contact note | 0.28 (1.78) | 0.00 (0.00) | 0.33 (2.01) | 0.00 (0.00) |
| Did not indicate next check-in date | 4.95 (9.24) | 3.82 (6.64) | 0.43 (2.15) | 1.93 (5.64) |
| Content audit | | | |
| Critical tone/lack of praise | 3.80 (7.49) | 3.43 (5.82) | 0.79 (3.63) | 0.29 (1.91) |
| Unresponsive to symptom increase, scoring measures, or suicide risk | 12.02 (19.10) | 9.56 (15.93) | 6.93 (14.29) | 4.45 (9.81) |
| Lack of psychoeducation | 5.03 (9.10) | 6.05 (12.39) | 1.75 (5.18) | 1.50 (6.39) |
| Does not answer patient question | 1.40 (4.71) | 0.77 (3.43) | 0.00 (0.00) | 1.17 (3.78) |
| Does not address patient concern | 10.30 (22.85) | 8.56 (16.29) | 3.49 (11.09) | 3.98 (6.37) |
| Failing to ask one question to patient | 4.20 (8.23) | 4.69 (7.87) | 5.28 (13.42) | 8.70 (15.03) |
| Poor writing | 4.12 (7.99) | 6.92 (12.71) | 5.21 (8.47) | 4.69 (8.07) |
| Poor timing | 2.23 (6.54) | 2.89 (7.47) | 0.78 (3.26) | 0.21 (1.39) |
| Poor instructions | 11.49 (17.12) | 7.61 (9.30) | 9.55 (14.15) | 7.65 (11.00) |
| Did not link to course content | 6.37 (11.15) | 6.56 (11.24) | 6.16 (10.30) | 6.01 (10.67) |
| Unnecessary self-disclosure | 0.26 (1.60) | 0.99 (3.55) | 0.00 (0.00) | 0.55 (2.52) |

Note. Each audit cycle is 3 months long with first audit cycle reported starting October 1, 2018 and last audit cycle ending September 30, 2019. Scores are presented as percentages of messages per patient containing each therapist behaviour.
from AC1 to AC4 and were below 5% at AC4. Four items that did not follow this pattern were Failing to Ask One Question to Patient, Poor Instructions, Did not Link to Course Content and Poor Writing. Failing to Ask One Question to Patient increased from 4.2% (SD = 8.23) in AC1 to 8.7% (SD = 15.03) in AC4. Poor Instructions decreased from 11.49% (SD = 17.12) in AC1 to 7.65% (SD = 11.00) in AC4 but was not below the targeted 5%. Did not Link to Course Content remained stable and above the 5% target across all audits (> 6.01%). Poor Writing was just below 5% at AC1 and AC4 but increased slightly in ACs 2 and 3.

3.2. Therapist example

To indicate how the audit can be examined individually, please see Table 2. For this example, one therapist was chosen to illustrate use of A&F at the individual level. As evident from Table 2, most of the percentage scores were under 5%, indicating a positive rating on most of the items during all ACs (e.g., during AC1 3/4 Basic Practice items, 8/11 content items). At AC1, Critical Tone, Poor Writing and Poor Instructions were above 5%; as described above, Poor Instructions was observed as being a common issue for the team, but Critical Tone and Poor Writing were unique to the individual therapist. At AC4, four behaviours were identified as above 5%. Poor Writing and Poor Instructions had improved but remained higher than 5% and Failing to Ask One Question to Patient and Did not Link to Course were not initially undesirable but increased above 5%. Poor Writing was unique to this therapist while the other behaviours were noted in the group analysis.

4. Discussion

In this paper, we have described the application of A&F as a quality improvement method for therapist-assisted ICBT. Our results indicate that quality could not only be assessed (i.e., quality assurance), but also improved in a continuous manner. Completion of four ACs created a sense of community and teamwork. The AC process was also deemed helpful for the therapists in improving care of patients. In conducting this A&F, we followed Colquhoun et al. (2017b) when formulating our A&F approach focusing on the six categories of whom, what, why, when, how, and how much. The list of modifiable therapist behaviours to audit was based on past research and clinical practice in our unit (Hadjistavropoulos et al., 2019a; Hadjistavropoulos et al., 2018b). In ICBT, there are as of yet no gold standards for therapist behaviours and as such our form was based on standards within our unit that are informed by the research literature. Although we discussed among our team the possibilities of adding, deleting, or modifying additional therapist behaviours, having an initial audit based on the literature as well as clinical practice in our unit provided a good starting point, especially because there remains relatively little research to draw on examining therapist behaviours in ICBT.

In all four audit periods, we found a low percentage of undesirable therapist behaviours (i.e., therapists displayed the behaviour in 12% or less of emails sent). For most of the therapist behaviours, we saw a trend towards improvement across the four ACs. Nevertheless, there were three behaviours that were low throughout all ACs (i.e., No Contact Note, Does not Answer Patient Question, and Unnecessary Self-disclosure) perhaps representing behaviours not requiring subsequent audit. Three items (i.e., Failing to Ask One Question to the Patient, Poor Instructions, and Not Linking Email to Course Content) were above the set target at the end of the four ACs and raised questions for our team about how important these behaviours are for patient care (e.g., patient engagement), the factors that may contribute to their elevation (e.g., workload) and methods of improving therapist scores (e.g., reminders).

Both the group level and the individual analyses were helpful for the therapists. Overall, the group level analysis was deemed valuable as it allowed therapists as a group to discuss and reflect on guidelines that were set and confirm commitment to practice standards. Bringing the areas that needed improvement to the therapists’ attention on a regular basis with objective indicators allowed for improved focus on the guidelines and allowed for regular discussion about whether these needed to be adjusted or whether practice could be changed. Doing so in full collaboration with the therapists allowed for a sense of community and teamwork. The AC process was also deemed helpful on an individual therapist level with feedback from therapists suggesting that they appreciated the more specific feedback they received in AC3 and AC4. As highlighted with the therapist case example, therapists could observe the individual therapist behaviours they had to work on.

There were a number of factors in this A&F that were based on our “best guess”, such as what behaviours to audit, what cut-off score to choose for problematic therapist behaviours (i.e., we chose 5%), what feedback to give, how many cases to audit, how often to provide feedback, or by what means the feedback should be given to therapists. Although we relied upon the literature to guide these decisions, we are unsure whether they were optimally derived or whether our results might have been different with another approach. Ultimately, part of the reason for engaging in A&F was to be able to demonstrate to our funders that we set and monitor (Titov et al., 2019).

Overall, A&F was instructive for our team not only for monitoring and improving undesirable therapist behaviours, but also for raising

| Table 2 | Descriptive statistics from one therapist showing mean percentage of times each therapist behaviour was present in messages audited by audit cycle. |
|---|---|
| | Audit cycle 1 | Audit cycle 2 | Audit cycle 3 | Audit cycle 4 |
| | M% (SD) | M% (SD) | M% (SD) | M% (SD) |
| N cases = 3 | N cases = 10 | N cases = 14 | N cases = 7 |
| Basic practice | | | | |
| 1. Did not message | 8.33 (7.22) | 2.11 (4.46) | 5.16 (1.027) | 0.00 (0.00) |
| 2. Did not call as indicated | 0.00 (0.00) | 1.11 (3.51) | 0.71 (2.67) | 3.37 (6.42) |
| 3. No contact note | 3.70 (6.42) | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) |
| 4. Did not indicate next check-in date | 3.70 (6.42) | 1.11 (3.51) | 1.54 (3.98) | 0.00 (0.00) |
| Content audit | | | | |
| 1. Critical tone/lack of praise | 7.87 (6.85) | 2.00 (4.21) | 0.71 (2.67) | 0.00 (0.00) |
| 2. Unresponsive to symptom increase, scoring measures, or suicide risk | 0.00 (0.00) | 5.33 (9.02) | 9.09 (12.30) | 3.21 (5.54) |
| 3. Lack of psychoeducation | 0.00 (0.00) | 0.00 (0.00) | 1.79 (6.68) | 3.17 (8.40) |
| 4. Does not answer patient question | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) |
| 5. Does not address patient concern | 0.00 (0.00) | 5.56 (12.00) | 0.65 (2.43) | 3.17 (5.42) |
| 6. Failing to ask one question to patient | 0.00 (0.00) | 2.22 (7.03) | 4.92 (10.15) | 16.27 (12.87) |
| 7. Poor writing | 15.74 (13.70) | 5.22 (7.61) | 8.81 (10.63) | 7.30 (8.29) |
| 8. Poor timing | 3.70 (6.42) | 1.00 (3.16) | 2.16 (4.31) | 0.00 (0.00) |
| 9. Poor instructions | 29.17 (26.02) | 6.72 (9.55) | 18.25 (15.01) | 11.51 (9.68) |
| 10. Did not link to course content | 4.17 (7.22) | 8.58 (8.41) | 4.96 (9.88) | 6.19 (8.65) |
| 11. Unnecessary self-disclosure | 0.00 (0.00) | 0.00 (0.00) | 0.00 (0.00) | 3.37 (5.77) |

Note. Each audit cycle is 3 months long with first audit cycle reported starting October 1, 2018 and last audit cycle ending September 30, 2019. Scores are presented as percentages of messages per patient containing each therapist behaviour.
questions about what we are willing to accept as appropriate and what we are not. This not only entails choosing our therapist behaviours, but also the percentages that we think should be addressed. We questioned whether having a blanket “cut-off” score (i.e., 5%) was appropriate given the different items on our audit measure. For instance, is 5% too low for deeming feedback necessary for behaviour such as Unresponsive to Symptom Increase, Scoring Measure, or Suicidal Risk compared to the presence of behaviour such as Poor Writing? Are we expecting too high percentage ratings from our therapists given the workload of a busy practice? Do these ratings impact patient outcomes? Are there areas that we should never accept error in and are there areas in which errors are more acceptable? If so, should these two areas perhaps be monitored more and less frequently? Given the cyclical nature of A&F, we question whether we should continue monitoring all these behaviours in subsequent ACs, or rather be more strategic in our focus.

4.1. Limitations and future directions

There are some important limitations that need to be mentioned. In terms of methodology, we did not assess inter-rater reliability between auditors. Further, we only audited 18% of patients; there was also some variability in the percentage of therapist caseloads that were audited, with some therapists having a greater percentage of their caseloads audited than others. It is also notable that during AC2 and AC3 a new therapist joined the team during these cycles, which has potential to impact audit scores. We did not examine relationships between audit scores and patient outcomes. It may be helpful to examine patient data to see how the audit scores relate to patient functioning. For example, we could determine whether the therapist audit scores are lower with certain patients (e.g., patients who ceased treatment early or patients who did not improve overall). Past research for example has shown that it is more difficult for therapists to follow practice guidelines when patients are less engaged in ICBT (e.g., Hadjistavropoulos et al., 2018b).

A&F is not a static process, but rather is dynamic, enabling the clinical team to alter the measurement of behaviours most fitting for facilitating concrete clinical changes (Pedersen et al., 2018a). In the future, we could change which behaviours we monitor, for example focusing somewhat less on undesirable therapist behaviours and more on desirable therapist behaviours. Of note, with the current audit, the decision was specifically made to focus on the presence of undesirable therapist behaviours or poor quality behaviours that could be improved rather than measuring the presence of desirable therapist behaviours as captured by another scale developed by our team, namely the ICBT Therapist Rating Scale (ICBT-TRS; Hadjistavropoulos et al., 2018b). Nine therapist behaviours are included on the ICBT-TRS capturing whether therapists exhibit expected behaviours in emails sent to patients, including builds rapport, seeks feedback, provides symptom feedback, provides psychoeducation, facilitates understanding, praises effort, encourages practice, clarifies administrative procedures, and communicates effectively (Hadjistavropoulos et al., 2018b). However, in the current study we chose to focus on undesirable behaviours as, in the literature, it has been emphasized that behaviours to be focused on are ones that require change (Colquhoun et al., 2017a). Although positive behaviours could also be changed (e.g., providing encouraging comments when this was not previously done), we deemed it more time efficient and critical to quality assurance for our funders to ensure undesirable behaviours were not present. Certainly, there is room for improvement in the audit process. As noted by one of our auditors, it was sometimes the case that therapists with a higher number of undesirable therapist behaviours, could still be sending patients very good messages. For example, the undesirable therapist behaviours (e.g., not linking to course materials, poor instructions) could be offset by many positive therapist behaviours (e.g., answering patient questions, empathy). There were also instances where an auditor felt that the audit did not identify an undesirable therapist behaviour, but still felt that the message was problematic in some way (e.g., message lacked warmth or was very generic). In the future, it might also be instructive to involve patients in the review of emails or selection of therapist behaviours of importance to monitor, which would be consistent with a patient-oriented approach.

In the future, we could also determine the appropriate timing for giving feedback (e.g., more often than every 3 months to facilitate more immediate change). In terms of the process of giving feedback, Ivers et al. (2014) state that this tends to be most helpful when accompanied by information that allows comparison of performance with a group of peers. To date, we have not provided therapists with specific details on percentage of undesirable therapist behaviours across all therapists; given the small number of therapists we were concerned that this specific comparative data could reduce morale. Instead of providing it to therapists, we are currently providing the details to supervisors. In the future, however, it may be beneficial to give the specific percentage of negative therapist behaviours across all therapists to each therapist. Pedersen et al. (2018a) further state that feedback followed by an action plan is most effective. It is possible that we could be more directive with therapists with respect to behaviours to work on and also engage in period of close supervision for behaviours of concern. Instead, as a group, we have emphasized that all therapists should work on behaviours. Changing how we give feedback so it is either accompanied by or followed by a directive, yet jointly constructed, action plan may ensure that the targeted therapist behaviours are corrected to a greater degree. An additional future direction would be to systematically use qualitative methodology to explore therapist experiences with the A&F (e.g., focus groups, structured interviews). Indirect feedback from therapists is that the A&F process is beneficial to their practice; nevertheless, we recognize that this feedback is informal and with a small group of therapists this feedback may be subject to biases.

It would be valuable to examine whether this A&F is general enough to be applied outside of our specific model of care to other ICBT routine care settings (e.g., Titov et al., 2018), or to other ICBT programs (e.g., to the Pain Course; Hadjistavropoulos et al., 2018a). Given that A&F developed for this study was based on literature (e.g., Colquhoun et al., 2017b; Ivers et al., 2012; Hysong et al., 2012), it is quite possible that all or some parts of the A&F form and process is applicable to other ICBT settings.

5. Conclusions

A&F is a process commonly used to ensure quality improvement among clinical teams in mental health care settings (e.g., Pedersen et al., 2018a). To our knowledge, it has not been previously used as a tool for monitoring and improving therapist-assisted ICBT in routine care. We found A&F to be helpful for identifying areas both in which our therapists were performing well in and areas in which they could improve. These areas were clarified both for individual therapists as well as for the team. We were able to provide feedback in multiple ways (i.e., individual supervision, group email, group supervision) depending on whether the undesirable behaviour was individual or group based. The A&F process raises interesting questions about the extent to which therapist behaviours are common across therapist-assisted ICBT programs in routine care and also about levels of acceptable deviation from practice.

Declaration of competing interest

All authors declare that they have no conflict of interest.

Acknowledgements

The authors would like to thank the research staff, therapists, and patients associated with the Online Therapy Unit and its community partner unit, the Saskatchewan Health Authority’s Online Therapy Team.

Funding

H.D.H. is funded by Saskatchewan Ministry of Health to operate the Online Therapy Unit in collaboration with the Saskatchewan Health Authority who is also provided Ministry of Health funding to hire online...
therapists. Research on the Wellbeing Course within the Online Therapy Unit is currently funded by the Canadian Institutes of Health Research (152917). Funders had no involvement in the design of the paper, collection, analysis, or interpretation of the data. N.T. and B.F.D. are funded by the Australian Government to operate the national MindSpot Clinic.

Appendix 1. Quality control audit

| Basic practice                  | Number of emails where undesirable therapist behavior was observed |
|--------------------------------|---------------------------------------------------------------|
| Did not message                | 0 1 2 3 4 5 6 7 8 9 10                                     |
| Did not call as indicated      | 0 1 2 3 4 5 6 7 8 9 10                                     |
| No contact note               | 0 1 2 3 4 5 6 7 8 9 10                                     |
| Did not indicate next check in date | 0 1 2 3 4 5 6 7 8 9 10                                   |

| Content audit                  | Number of emails where undesirable therapist behavior was observed |
|--------------------------------|-------------------------------------------------------------------|
| Critical tone/lack of praise   | 0 1 2 3 4 5 6 7 8 9 10                                           |
| Unresponsive to symptom increase, scoring measures, or suicide risk | 0 1 2 3 4 5 6 7 8 9 10                                         |
| Lack of psychoeducation        | 0 1 2 3 4 5 6 7 8 9 10                                           |
| Does not answer patient question | 0 1 2 3 4 5 6 7 8 9 10                                     |
| Does not address patient concern | 0 1 2 3 4 5 6 7 8 9 10                                      |
| Failing to ask one question to patient | 0 1 2 3 4 5 6 7 8 9 10                                     |
| Poor writing                   | 0 1 2 3 4 5 6 7 8 9 10                                           |
| Poor timing                    | 0 1 2 3 4 5 6 7 8 9 10                                           |
| Poor instructions              | 0 1 2 3 4 5 6 7 8 9 10                                           |
| Did not link to course content | 0 1 2 3 4 5 6 7 8 9 10                                           |
| Unnecessary self-disclosure    | 0 1 2 3 4 5 6 7 8 9 10                                           |

Content Audit Score (total number of emails with each undesirable therapist behaviour):

Comments:

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________
Were any coaching topics identified? Yes/No

Type of coaching:
- Individual coaching
- Group reminder message
- Group supervision session

Notes:________________________________________________________________________
______________________________________________________________________________
______________________________________________________________________________
______________________________________________________________________________
______________________________________________________________________________
______________________________________________________________________________

This case would make a good training case: Yes/No

Audit form definitions
Did not message: Did not write to a patient when it was expected (e.g., regular check-in day).

Did not call as indicated: Did not call patient when there was a request to call, suicide risk, hadn't logged in for over a week, or when symptoms scores increased by 5 points or more.

No contact note: A written contact note was not added to the patient's website file to indicate contact outside of standard messaging or to add a concern/update to the file.

Did not indicate next check-in date: Next expected contact date was not included in the message to the patient.

Critical tone/lack of praise: The writer's message could be considered negative, scolding, abrupt or critical in terms of language or writing style, and/or no supportive comments, praise, or encouragement.

Unresponsive to symptom increase, scoring measures, or suicide risk: Does not provide symptom feedback, inquire about symptoms, or does not respond to increased risk.

Lack of psychoeducation: The writer did not provide at least some minimal information about lesson(s), skills, resources available and/or information about the nature of depression and anxiety (or addressed the wrong lesson) if applicable.

Does not address patient concern: The writer does not follow up on, or fails to address, the patient's specific concerns or feedback, and/or expression of current issues when responding to the patient. The message is not customized to the patient and their concerns.

Does not answer patient question: The writer does respond to a patient's specific question.

Failing to ask one question: The writer does not ask one question or inquire how the patient is functioning, or ask if the patient has any concerns or questions.

Poor writing: The message is poorly written with respect to style, spelling/grammar, language, or length (2 or more spelling/grammar issues or one more major issue that could cause confusion.)

Poor timing: The writer did not address a patient concern in a timely matter, shared information about a lesson/resource and was off with respect to timing.

Poor instructions: The writer did not offer adequate instructions about what is expected at this time, what's coming up, or what is required in the future.

Did not link to course content: The writer did not make connections between patient experience and course content.

Unnecessary self-disclosure: Writer's self-disclosure detracted from the patient's therapeutic experience.

References
Anderson, G., 2016. Internet-delivered psychological treatments. Annu. Rev. Clin. Psychol. 12, 157–179. https://doi.org/10.1146/annurev-clinpsy-021815-093006.
Anderson, G., Carlbring, P., Titov, N., Lindefors, N., 2019. Internet interventions for adults with anxiety and mood disorders: a narrative umbrella review of recent meta- analyses. Can. J. Psychiatr. 64, 465–470. https://doi.org/10.1177/0706747719839381.
Aveline, M., Watson, J., 2001. Making a success of your psychotherapy service: the contribution of clinical audit. In: Mace, C., Moorey, S., Roberts, B. (Eds.), Evidence in the Psychological Therapies: A Critical Guide for Practitioners. Brunner-Routledge, Hove, East Sussex, pp. 199–210.
Carlbring, P., Anderson, G., Cuijpers, P., Riper, H., Hedman-Lagerlöf, E., 2018. Internet-
