How to Build a Research Database from Data You Collect to Guide Your Clinical Work

Jacqueline B. Persons,[a, b] Cannon Thomas,[c, d] Megan Hsiao[a] and Rebecca Courry[a]

[a]Oakland Cognitive Behavior Therapy Center

[b]University of California, Berkeley

[c]San Francisco Group for Evidence-Based Psychotherapy

[d]University of California, San Francisco

In press, *Cognitive and Behavioral Practice*

We thank the patients and therapists who contributed their data, and Connie Fee, Garret G. Zieve, Vael Gates, Emily Nicole Bailey, and many volunteer research assistants, especially Corey Pallatto Hughan, for their help with the database described here. Parts of this article were presented by Cannon Thomas at the annual conference of the Association for Behavioral and Cognitive Therapies in Philadelphia in 2014.

Correspondence concerning this article should be addressed to Jacqueline B. Persons, Oakland Cognitive Behavior Therapy Center, 5625 College Avenue, Suite 215, Oakland, CA, 94618. E-mail: persons@oaklandcbt.com

Abstract

Clinicians who collect progress monitoring data from their clients to guide their clinical work are accruing valuable data that can contribute to the research literature. Our goal in this article is to help clinicians create a research database from the data they collect to guide their clinical work. We begin by offering a rationale to clinicians for undertaking the effort to create a research database. We describe our own experience creating a research database based on our clinical data, and we describe the published empirical studies based on the database. In the main body of the article, we offer step by step guidance to clinicians who wish to create their own research database. We list thirteen steps, beginning with "Select a measure or measures you want to use for progress monitoring . . ." that clinicians can take to build a research database from the data they collect to guide their clinical work. Finally, we offer a brief description of resources that can help clinician-researchers with some of the other elements of the research process, including obtaining library access and handling ethical issues.

Keywords: practice-based research, private practice, measurement-based care

How to Build a Research Database from Data You Collect to Guide Your Clinical Work

When clinicians who rely on evidence-based practices (that is, they are *consumers* of science) collect progress monitoring data to guide their work, they can aggregate these data into a database that can serve research purposes, and these same clinicians can become *producers* of science. Clinicians may feel uncertain how to proceed to create a research database, and they may feel unconvinced the task is worth the trouble. We offer guidance in this article. We highlight the advantages to clinicians of creating a research database. We describe a research database we created from data collected in routine clinical practice, and the empirical studies we published using data from the database. Our major goal in this paper is to provide step by step guidance to clinicians who wish to create their own research database from data they collect for clinical purposes. Of course, collecting the data is only one element of the research process. We conclude our paper with a brief description of resources to help the clinician with other aspects of the research enterprise that are outside the scope of this paper. We work in a private practice setting, and the methods we describe arise from our experiences in that setting; however, we expect them to be useful to providers working in other clinical settings.

Why do it?

Creating a research database from the data you collect for clinical purposes is worth doing for several reasons. First, the effort to create a research database can increase your motivation to collect progress monitoring data to guide your clinical work. Progress monitoring is an essential element of evidence-based practice (American Psychological Association, 2005). Progress monitoring has been shown to lead to better patient outcomes; a recent review paper (Lewis et al., 2018) showed that patients treated by clinicians who use measurement based care (defined as "the systematic evaluation of patient symptoms before or during each clinical encounter to inform behavioral health treatment" (p. E2) have better outcomes than patients whose clinicians do not do this. For example, as part of a discussion with the client about her lack of progress, the last author reviewed with her client the plot of her Depression scores on the Depression Anxiety Stress Scores (DASS; Lovibond & Lovibond, 1995). Because the therapist was collecting data at every session (rather than monthly or on an ad hoc basis), she and her client, when viewing the plot of the client's data, were able to identify a reliable increase in symptoms that occurred every month. Further investigation revealed that this monthly blip was tied to the client's menstrual cycle. Armed with this information, the client consulted with her pharmacotherapist, who made some adjustments in the client's medication that led to a large reduction in her symptoms.

Second, if you aggregate your data into a database, you can use it to learn things that will help you improve the quality of the care you provide. For example, you can find out whether your outcomes are similar to those of comparable patients treated in the randomized controlled trials (Persons, Bostrom, & Bertagnolli, 1999) and whether your outcomes vary as a function of the problem domain of the patient you are treating (Kraus et al., 2016; Boswell, Constantino, Kraus, & Coyne, 2020). Learning that you have better or worse outcomes with some problem domains can help you select which patients you want to treat, and what additional training you might wish to obtain.

Third, creating a database gives you the opportunity to influence the field and contribute to science (Kazdin, 2008). You can contribute to studies of research questions that flow out of your clinical experience and your unique vantage point as a clinician. For example, the first author's clinical experiences led her to the hypothesis that the case formulation is a useful tool for solving clinical problems, and systematic data collection in her practice allowed her to publish empirical papers to test that hypothesis. She published an open trial showing that depressed outpatients treated with case formulation-driven CBT for depression had outcomes comparable to those of patients who received CBT guided by a standardized protocol in randomized controlled trials (Persons, Roberts, Zalecki, & Brechwald, 2006) and a paper showing that a written case formulation in the clinical record was associated with reduced rates of dropout (Gates, Hsiao, Zieve, Courry, & Persons, 2021). Think about what you have learned from your clinical experience that you believe could contribute to knowledge in the field, and use that idea to guide your data collection.

Fourth, building a research database can help you build a network of colleagues that can provide you with a rewarding antidote to the isolation that many practitioners, especially those in solo private practice, experience. As a clinician-researcher you can join the community of scientists and scholars as you work with assistants and collaborators and journal editors and reviewers on your projects, and present them at national or, if you like to travel, international conferences. The ABCT (and its Special Interest Group for Research in Clinical Practice) provides an excellent home for clinician-scientists, as do several other professional societies, including the Society for a Science of Clinical Psychology (SSCP) and the Society for Psychotherapy Research.

Fifth, building a research database allows you to contribute to reducing the oft-described gap between science and practice, fulfilling the ideal of the scientist practitioner described in the Boulder model of clinical psychology (Benjamin Jr & Baker, 2000).

Finally, creating a research database helps clinicians meet the obligations laid out by their professions' codes of ethics. For example, both the American Psychological Association (2017) and the American Association for Marriage and Family Therapy (2015) state that members should devote a portion of their efforts to activities that benefit the community for which there is little or no financial return. Clinicians often meet this obligation by providing sliding scale direct services. Building a database and contributing to research is another way to meet this ethical obligation. In fact, the code of ethics of the National Association of Social Workers (2008) specifically calls for clinicians to "contribute to the profession's literature and to share their knowledge at professional meetings and conferences."

Our Experience Building a Research Database from Data Collected for Clinical Purposes

The first author was trained (by Aaron T. Beck) to ask her depressed patients to complete a standardized symptom scale at the beginning of every therapy session, to plot the scores on a graph, to review the plot with the patient in the session, and to use the data to guide decision-making. She was trained in a university medical center (that is, a research center), where another of her teachers (David D. Burns), modelled the practice of conducting research in his clinical practice and taught her to include in her treatment agreement a section asking the

patient to give permission for use of the clinical data for research purposes. From the years 1981 to 2009, she collected these symptom data and consent documents and stored them in her paper clinical records. During part of that time, she worked in a group practice with nineteen other clinicians who were part of the practice for varying lengths of time. Many were her trainees, and others were colleagues who adopted her approach to collecting symptom data and consent for research. As a result, after many years of data collection, the clinicians in the practice had accrued symptom tracking data from nearly 1500 patients.

These data were gathered for clinical purposes. At the same time, the first author, dating from her training experiences, anticipated the potential of conducting research based on the clinical data, and for that reason always included in her first treatment agreement a section in which she asked her patients for permission to use their data for research purposes. Without this informed consent document, it would not have been possible to create a research database from the clinical data. Over time, the practice developed other procedures that supported both good clinical work and research goals. For example, in 2001 the practice developed a standard packet for gathering demographic and background information that was completed prior to intake by all patients from that date onward.

The data remained in the (paper) clinical files until 2007, when the first two authors, aided by many volunteer research assistants and by the clinicians in the practice, undertook the laborious process of culling the data from the files and integrating it into the Naturalistic CBT Archival Database. Every patient aged 18 and over who had given consent for use of their data for research purposes and who had completed at least one Beck Depression Inventory (BDI) or Burns Anxiety Inventory (Burns AI) was included in the database. Because the BDI and Burns AI were administered to every client prior to intake and were the core outcome measures used for tracking progress, the database included the vast majority of cases treated by the first author and her colleagues over those 27 years: 1470 cases in all. We included in the database item-level data for every standardized measure administered to each client. In addition to the primary measures, we included item-by-item scores on 10 additional measures we found in the clinical files (e.g., the Yale-Brown Obsessive-Compulsive Scale). The BDI was by far the most common measure used for symptom tracking, with a total of 18,602 administrations. In addition to the symptom measures, we included in the database demographic and background information for every patient (trauma history, current substance use and abuse history, family history, psychiatric history, etc.). For the roughly two thirds of patients who had completed a standard intake packet, we used their responses to the standard questions to complete this information. Otherwise, the information was gleaned from the written file wherever possible. We also recorded the psychiatric diagnosis given by the clinician at intake and termination, whether certain decision support tools were present in the file (a written case formulation, a list of treatment goals, and a plot of progress monitoring data with at least one score on it), and a few therapist-rated judgments about the nature of the termination and the reasons for termination.

The process of pulling these data from the clinical records and entering them into the database required many hundred person-hours of work. But the resulting database is large and unique and has contributed (at the time of this writing) to the publication of nine empirical papers, with one more project underway (see Appendix). The papers have made useful contributions to the study of the effectiveness of naturalistic CBT and case formulation-driven CBT; the predictors of dropout and outcome of naturalistic CBT; and the trajectory of change,

including sudden gains in CBT. Of course, the process of collecting data is only one part of the work needed to produce publishable research. At the end of this paper, we point readers to resources to help with other aspects of the research process. We focus here on the details involved in creating a database.

Our experience creating the Naturalistic CBT Archival Database has taught us a great deal about how to make the process of accruing clinical data for scientific purposes more effective and efficient. With the assistance of modern technology and some forethought, we believe the task of constructing a research database from data that were collected for clinical purposes can be accomplished more effectively with a fraction of the time and effort we spent. In the next section we offer our recommendations to clinicians who wish to use the data they collect for clinical purposes to create a research database.

How to Create a Research Database in a Clinical Practice Setting

We list here in approximate order of importance the steps required to create a research database from the data you collect to guide your clinical work. You can do the work gradually. Just take one step at a time, and if you have colleagues in your group practice or organization, divide up the work.

1. Select a measure or measures that you want to use for progress monitoring, and an online tool to administer them. The single most important lesson we learned from our work to create the Archival Database is that an online tool is needed so that patients can enter their data, thus saving the clinician the work of entering it, and reducing the potential for errors. We recommend selecting an online tool that has these features: HIPAA-compliant; access to a library of symptom measures that are clinically-useful and widely used by researchers; patients can enter data on their smartphones or browser; the tool creates a plot of the scores; data can be downloaded by the therapist in comma separated values (CSV) format or a similar format that can easily be exported to a software tool for statistical analyses; affordable; good customer support. The tool we've discovered that has all these features and have used successfully for about 18 months is PsychSurveys (www.psychsurveys.com).

We recommend that clinicians who plan to create a research database select one measure that all or nearly all the patients in the practice will complete, in order to create a database suitable for group-based studies. Persons and Courry at the Oakland CBT Center selected the twenty-one item Depression Anxiety Stress Subscale (DASS; Lovibond & Lovibond, 1995) for several reasons. First, its three subscales were based on Clark & Watson's (1991) tripartite theory and thus permit assessment of negative affect (the Stress scale), low positive affect (the Depression Scale) and physiological arousal (the Anxiety scale) that the tripartite theory viewed as independent; the strategy of assessing three relatively independent sets of symptoms in one measure was appealing to us. Second, the scales assess symptoms of anxiety and depression, which are the most common symptoms encountered at our Center and in most private or community practice settings. Third, the scales are responsive to change during treatment. Fourth, the measure has been shown in several published studies to have good psychometric properties (e. g., Antony, Bieling, Cox, Enns, & Swinson, 1998; Brown, Chorpita, Korotitsch, & Barlow, 1997). Finally, the measure is in the public domain, and is available free of cost. A disadvantage of the DASS-21 for routine outcome monitoring is that it does not assess suicide. We added two

suicide items to the DASS-21, creating what we call the DASS-23 (the measure is posted on our website at https://oaklandcbt.com/forms-and-tools-for-clinicians).

In addition to the DASS-23, the therapists at the Oakland CBT Center also use the Patient Health Questionnaire-9 (PHQ-9; Kroenke, Spitzer, & Williams, 2001), which assesses depressive symptoms, and the GAD-7 (Spitzer, Kroenke, WIlliams, & Löwe, 2006), which assesses symptoms of Generalized Anxiety Disorder and anxiety symptoms more generally. The PHQ-9 has a 10th item that assesses the degree to which the symptoms interfere with functioning. The PHQ-9 and GAD-7 are in the public domain and available at no cost and can be downloaded at https://www.phqscreeners.com/, where a drop-down menu provides access to the measures in dozens of languages. Other clinically-useful measures in a typical outpatient population include the Yale-Brown Obsessive Compulsive Inventory (Goodman et al., 1989), and the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5; Blevins, Weathers, Davis, Witte, & Domino, 2015).

We strive to collect symptom progress monitoring measures at every therapy session. From a clinical point of view, data collection at every session is important because even when the patient's scores on the DASS or the PHQ-9 are in the normal range, ongoing assessment can alert you to flareups. From a research point of view, collecting measures at every session is important to reduce biases that occur if measures are not missing at random, such as when the clinician's decision to collect the measure is based on the patient's clinical status.

One weakness of the Archival Database is that there is a substantial amount of missing data, and the missing data cannot be assumed to be random. The database has many cases where measures were given at the beginning of treatment, then were not given for several sessions, and then were given again. The decision to give or not give the measures in these cases may well have been driven by clinical considerations. For example, a clinician might be less likely to give measures when patients are less symptomatic, and then resume giving the measure when symptoms flare up. This type of decision-making about administering the measures can affect the results of research relying on the database. In this example, data analyses would show less improvement than actually occurred, because measures were administered only when symptoms were more severe. To prevent these sorts of biases, we strongly recommend developing a clear protocol for gathering data at pre-determined intervals (every session or monthly, for example) and developing procedures to maximize adherence. If a measure cannot be completed at a particular session, it is important to record the reason the data are missing, in order to promote clinician adherence to progress monitoring and help interpret the missing data when writing up a research study based on the data.

2. Build data collection into your routine clinical practice.

Building a research database in a clinical practice setting is feasible only if data collection is an essential element of the clinical work. At the Oakland CBT Center, we collect one or two or three standardized symptom monitoring measures at every session for almost every patient. We also collect a standard set of intake measures (described below, in item 5), and rely on a standard progress note (described below, in item 6), and a standard feedback tool at every session (described below, in step 6).

In our experience, patients rarely balk at data collection. They are generally eager to save time during the assessment interview by completing measures before the interview, and to provide data to monitor their progress in therapy. When clients do exhibit ambivalence about completing measures, the clinician can increase the client's motivation to provide the data by reviewing the client's scores on the measures with the client in the session and using the data to guide the treatment, so the client sees that the clinician is actually using the data the client provides. Providing a user-friendly online tool (see item 1 in this list) also promotes compliance with data collection procedures. And when clinicians learn how useful the data can be to demonstrate their effectiveness and guide their work, data collection becomes a helpful part of the clinical work rather than a tiresome obligation.

3. Request consent for research in your standard Treatment Agreement.

International codes of ethics view informed consent as an essential element of ethical research (Amdur & Bankert, 2010), and those ethical standards have been adopted by our professional associations. The elements of informed consent listed in the APA Ethical Code (section 8.02; American Psychological Association, 2017) are: the purposes of the research, expected duration, and procedures; the right to decline to participate and to withdraw; any consequences of declining or withdrawing; factors that may influence the client's decision to participate, such as risks or adverse effects; any potential benefits of participating; limits of confidentiality; incentives to participate; whom to contact to ask any questions about the research, and participants' rights. At the Oakland CBT Center, we recently re-wrote the section of our treatment agreement in which we ask for research so that it provides all of the elements required by the APA Ethical Code (section 8.02; American Psychological Association, 2017) and the NASW Code of Ethics (National Association of Social Workers [NASW], 2017), and we have posted it at https://oaklandcbt.com/resources-for-clinician-researchers for readers to access.

4. If your practice is a HIPAA-covered entity, add a HIPAA authorization for research to your intake documents.

The Oakland CBT Center developed a HIPAA authorization for research, and we post it at https://oaklandcbt.com/forms-and-tools-for-clinicians.

5. Use a standard intake packet for all your patients.

Think about what information you want at intake and ask all your patients to provide that information in writing before your intake interview. At the Oakland CBT Center our intake packet includes an intake questionnaire that collects a lot of information about the patient's demographics, personal history and current circumstances, medical and psychiatric treatment history, information about current and past use of substances, a diagnostic screening measure, and several standardized scales (the DASS-23, the PHQ-9, the GAD-7, the Work and Social Adjustment Scale (Mundt, Marks, Shear, & Greist, 2002), and the Obsessive Beliefs Questionnaire (OBQ-44; Obsessive Compulsive Cognitions Working Group, 2003)). All of these measures except the Work and Social Adjustment Scale are in the public domain, and we post them on our website at https://oaklandcbt.com/forms-and-tools-for-clinicians, and you are welcome to download and use them. We described above the research benefits of gathering information in a standard intake packet that is administered to all patients. Collecting a lot of basic information about the patient before the intake is helpful clinically as well. The first author

learned this lesson the hard way when, 10 minutes before the end of the initial assessment interview with a new patient, she learned that the man she was interviewing had a very serious substance abuse problem that she did not have the expertise to treat. If she had had that information at the beginning of the interview, she would have handled the interview much differently and been much more helpful to the patient. Each of the clinician-authors of this paper has had innumerable experiences where substance abuse issues were overlooked until much later in the treatment than was necessary, resulting in needless suffering.

We recommend that clinicians consider establishing some standard procedures for collecting diagnostic information at intake. Our papers based on the Archival Database have sometimes been criticized by peer reviewers because the procedures used to establish DSM diagnoses were not standardized. Diagnostic information is helpful clinically as well as for research; Youngstrom, Choukas-Bradley, Calhoun, & Jensen-Doss (2015) provided a persuasive argument that accurate diagnosis is the foundation for effective clinical work. The clinician might conduct a standardized diagnostic interview, identify a few specific questions to be asked during each intake, or use a diagnostic screening tool (for example, go to www.oaklandcbt.com/training/forms and tools for clinicians, where we post our Diagnostic Screening Tool; you are welcome to download it and use it). The answers to standard diagnostic questions could be recorded and added to a research database. As one example of the utility of this procedure, consider the difference a few specific questions about symptom history could make in the treatment of a depressed patient who has a history of previous hypomanic or manic symptoms. The first and second authors have had the experience of learning later that they missed a diagnosis of bipolar disorder in some of their intake evaluations, an error that resulted in substantial suffering for their patients.

When constructing your intake packet, be sure to include your standard progress monitoring measures (see item 1 above). Unless you collect pre-treatment data, you will not be able to fully and accurately measure change during treatment, because quite a lot of symptom change often happens very early in treatment. Many investigators have shown that early change (in symptoms of depression and anxiety and eating behaviors) predicts treatment outcome (e. g., Grilo, White, Masheb, & Gueorguiera, 2015; Vittengl, Clark, Thase, & Jarrett, 2019). In fact, one of the studies we conducted using our Archival Database showed that patients who remain severely depressed after 4 weeks of treatment are very unlikely to remit (Persons & Thomas, 2019).

6. Use a standard progress note and feedback measure.

A standard progress note format allows the clinician to collect the same information in every session and makes it easy to locate the information. At the Oakland CBT Center, the first author has begun using a Google Form for a progress note so the data can easily be exported to an excel document. We have not yet located an electronic health record that can support the easy and secure transfer of progress note data.

We also use a feedback measure, the Session Assignment and Feedback Form (SAFF; Persons, Hong, Lemle Beckner, Owen, & Eidelman, 2012), which asks patients to complete to give us information about their compliance with homework, skills they practiced, concepts they learned in the session, and the quality of the alliance. PsychSurveys recently developed an online version of this measure that you can ask for if you want to use it, or you can access our paper and

pencil version at www.oaklandcbt.com/training/forms and tools for clinicians. The data provided on the SAFF recently led to a creative study (Jensen et al., 2020) showing that homework compliance was better when therapists assigned homework related to the content of the patient's main take-home message than when the homework assignment was unrelated to the patient's take-home message.

A critique of some of our research papers based on the Archival database has been that we do not have much information about what happened in the therapy sessions. Creating a standard progress note and feedback measure provide a remedy for this problem. If you hypothesize that a particular intervention or procedure will have an impact on compliance, outcome, or dropout in your practice, using a simple "yes" or "no" checkbox makes it possible to test your hypothesis. Such an approach can be the engine behind innumerable contributions to the research literature.

Again, we note that the same procedures that produce a great research database also contribute to excellent clinical work. Having a standard progress note template makes documenting sessions easier and ensures that the most important information is recorded. For example, by recording whether you performed procedures you believe to be of particular importance to high quality therapy (e.g., setting an agenda for the therapy session), you hold yourself accountable to your own highest standards. Following a standard template, or "checklist," of best practices in this way has been shown to improve performance in a wide variety of fields (Gawande, 2009).

7. Think about what research questions are of interest to you as you design your data collection tools.

If you collect systematic data to guide your clinical work, you can amass a database that is useful for research purposes. That is because there is a very large overlap of questions clinicians want to know the answer to and questions researchers want to know the answer to (Persons, 2007). That said, we recommend thinking in advance about the questions that are most interesting to you so you can collect the data you need to study those questions. Do you want to measure your own outcomes over time to see if your skills improve? Test the effects of certain interventions? Learn more about the role of the therapeutic alliance in treatment? It's a good idea to sit down to write down the hypotheses you're interested in so you can plan your data collection accordingly.

At the Oakland CBT Center, as a result of an early collaboration with Amy Sanchez when she was a graduate student at UC Berkeley, we are interested in the relationship between skills use and outcome. We are also interested in the relationship between the quality of the alliance and symptom change. To tackle both of those questions, we ask our patients to report every week (on the SAFF, described above) what skills they used during the week, and the quality of the alliance in the previous session.

8. Document your data collection decisions and practices.

We recommend two pieces of documentation: Clinical Record Tools, and Database Documentation. The Clinical Record Tools document is a running log of what intake and progress monitoring, progress note, and other measures and software tools and procedures you use in your clinical practice, and where you store this information. Write down each change you

make, the date you make it, and your rationale for making the change. In the Database Documentation document, you can keep a running log of all the steps you take and the dates you take them to establish and maintain your database collection procedures.

9. Create a Participant List.

We recommend creating a list of patient participants that is separate from the progress monitoring data (see item 10). The Participant List is an Excel document that lists each patient's name, identification number, age at the first session, gender, race, marital status, date of first session (a separate column for month, date, year), and whether the patient gave consent for research. We also include (but only for patients who gave consent to provide data for research) in the Participant List certain other information we're interested in collecting to test our research hypotheses, such as (in our case) whether the patient terminated treatment uncollaboratively or prematurely, and whether the clinical record includes a written case formulation and a list of treatment goals. The Participant List includes protected health information, so you will want to handle it carefully. The Participant List is maintained by the therapist, and data from the Participant List is passed to the research coordinator only for patients who gave informed consent for research participation. If your practice includes more than one therapist, it is helpful if each therapist in the practice uses the same format for the Participant List so the lists can be merged (after names are omitted from the merged document). Include every patient who comes into the therapist's office, so that when you write the Method section of your paper, you'll be able to report the numbers of patients who do and do not give consent for research, and to evaluate whether patients who do and do not agree to research differ in basic demographics. In a recent calculation, we found that 112 of 124 patients (90.3%) treated by the first author consented to the use in research of data from their clinical record.

10. Build a database of your progress monitoring data.

The DASS Database at the Oakland CBT Center consists of an Excel document that includes item-by-item scores on the DASS-23 for each therapy session, where the first column of the row is the participant ID number, and subsequent columns include the month, day, and year of the session (a separate column for each element of the date), and the next 23 columns are the scores on the 23 items of the DASS for that patient for that session. Many of our patients return for more than one course of treatment, so our database has a column that indicates the number of the course of treatment that the session corresponds to. We define a course of treatment as ending when there is a termination note in the clinical record or 6 months or more passes since the previous therapy session. Only data from patients who gave written consent to provide data for research are entered in the database. We (or our research assistants) download data from our online progress monitoring tool into the database on a regular basis.

11. Clean the database.

Once the database is up and running, it is important to establish a plan for cleaning the database, either on a periodic basis, say quarterly, or when getting ready to analyze the data for a study. Simple spot checks can help with this process, and calculations can be conducted to search for out-of-range values, e.g., a score other than 0, 1, 2, or 3 on a DASS item.

12. Store your data securely. Back it up.

It is important to establish standards for security of physical and electronic data used by clinicians participating in data collection. Issues to consider include the physical security of the flashdrive or computer where data are stored, the security of passwords and cloud storage, and a review process to ensure that weak points in security are identified and addressed. We suggest creating a shared Data Security document that includes standards agreed upon by all parties. Questions to ask when developing your security policies include: What are the access points for data? What are our standards for password strength? How often are we updating our passwords? Are hard drives kept in locked cabinets? Are hard drives encrypted? Who has access to files and keys? How would we handle a laptop or phone theft? How is access removed when an assistant or clinician leaves a group?

To protect data, establish clear procedures for backing it up and for clearly communicating which version of the data is the most recent one. Whenever we update our database, we give the database a new name that includes the date the database was updated. Also, make sure that the backup plan conforms to the aforementioned security standards.

13. Automate processes whenever possible.

We recommend you automate some of the processes described above. For example, at the Oakland CBT Center we use a calendar application to automatically prompt security and document reviews at regular intervals. Our progress monitoring software tool can be programmed to prompt us to follow up if a measure is not completed by the client. We also recommend that you regularly evaluate how much time you spend managing your research database. Building sustainable practices is key to conducting projects that can take months or years to complete.

Other Elements of the Research Enterprise

We described steps clinicians can follow to build a research database from the data they collect for clinical purposes. However, conducting research entails much more than collecting data. We briefly point here to resources to aid the clinician-researcher with other aspects of the research process.

Attention to ethical and legal issues is an especially important element of the research process, and the integration of research and clinical practice poses particularly challenging issues. To address these issues, clinicians who are conducting program evaluation and research can benefit from consulting with colleagues, especially researchers and ethics specialists, and can obtain a review of their projects by an institutional review board or other group. A full discussion of this issue is outside the scope of this article; Persons et al. (2021) provide helpful guidance in handling the legal and ethical issues confronted by the practitioner who is conducting research in a clinical setting.

LeCompte and Young (2021) describe changes in the consent process and form that were implemented as a part of the 2017 revisions to the Federal Policy for the Protection of Human Subjects (the Common Rule). Wilson, Persons, Reiser, Osborne, & Rizvi (2019) offer suggestions to the clinician (and researcher) to help with the task of keeping up to date with the

research literature. Bem (2004) offers a primer for writing an empirical journal article. Appelbaum et al. (2019) describe journal article reporting standards for quantitative research, a report of the APA Publications and Communications Board Task Force. Kimball & Caserta (2004) provide a useful account of elegant strategies and guiding principles to attend to when organizing your database. Other topics, including selecting a good question, writing up and submitting the manuscript, finding collaborators and assistants, analyzing the data, are discussed in Persons (2021), Codd III (2018), and Castonguay, Barkham, Lutz, & McAleavey (2013). Useful descriptions of their business structure and infrastructure for conducting research in clinical practice settings are provided by LeJeune & Luoma (2015), and Osborne (2018). The ABCT Special Interest Group on Research in Clinical Practice can also be a useful resource to readers of this ABCT journal.

Concluding Remarks

Our goal has been to convince clinician readers that creating a research database using your clinical data is a worthwhile enterprise. We listed steps you can take to accomplish that task. Although our list includes thirteen steps, each one is straightforward and achievable. If you implement them, the result can be a gorgeous database that can contribute to knowledge in clinical science.

We wish you success in your data collection. Remember that if you have established a process for aggregating your clinical data into a research database, you are, every day, in every therapy session you conduct, making double the contribution. You are providing high quality care to your clients, and you are simultaneously contributing to science. What could be better than that?

References

Amdur, R. J., & Bankert, E. A. (2010). *Institutional review board: Member handbook*. Jones & Bartlett Publishers.

American Association for Marriage and Family Therapy. (2015). *AAMFT code of ethics*.

American Psychological Association. (2005). *American Psychological Association policy statement on evidence-based practice in psychology*. http://www.apapracticecentral.org/ce/courses/ebpstatement.pdf

American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct*. American Psychological Association. https://doi.org/10.1037/0003-066X.57.12.1060

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M.W., & Swinson, R. P. 1998. Psychometric properties of the 42-Item and 21-Item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment* 10(2): 176–81.

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

Attkisson, C. C., & Greenfield, T. K. (1994). Client Satisfaction Questionnaire-8 and Service Satisfaction Scale-30. In *The use of psychological testing for treatment planning and outcome assessment* (pp. 402–420). Lawrence Erlbaum Associates, Inc.

Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Bem, D. J. (2004). Writing the empirical journal article. In J. M. Darley, M. P. Zanna, & H. L. Roediger III (Eds.), *The compleat academic: A career guide* (Second edition), pp. 185–219). New York: Random House.

Benjamin Jr, L. T., & Baker, D. B. (2000). Boulder at 50: Introduction to the section. *American Psychologist*, *55*(2), 233.

Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., & Domino, J. L. (2015). The Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of Traumatic Stress* 28(6): 489–98.

Boswell, J. F., Constantino, M. J., Kraus, D. R., & Coyne, A. E. (2020). *Enhancing mental health care by matching patients to providers' empirically derived strengths* . Paper presented at the Association for Behavioral and Cognitive Therapies.

Brown, T. A., Chorpita, B. F., Korotitsch, W., & Barlow, D. H. (1997). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. *Behaviour Research and Therapy* 35(1): 79–89.

Burns, D. D., & Eidelson, R. J. (1998). Why are depression and anxiety correlated? A test of the tripartite model. *Journal of Consulting and Clinical Psychology*, *66*(3), 461.

Castonguay, L., Barkham, M., Lutz, W., & McAleavey, A. (2013). Practice-oriented research: Approaches and applications. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed.). Wiley.

Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: Evidence and taxonomic implications. *Journal of Abnormal Psychology*, *100*, 316–336. https://doi.org/10.1037/0021-843X.100.3.316

Codd III, R. T. (Ed.). (2018). *Practice-based research: A guide for clinicians*. New York: Routledge.

Davidson, J., Martinez, K. A., & Thomas, C. (2006). *Validation of a new measure of functioning and satisfaction for use in outpatient clinical practice*. Association for Behavioral and Cognitive Therapies, Chicago, IL.

Gates, V., Hsiao, M., Zieve, G. G., Courry, R., & Persons, J. B. (2021). Relationship to CBT outcome and dropout of decision support tools of the written case formulation, list of treatment goals, and plot of symptom scores. *Behaviour Research and Therapy*, *142*. https://doi.org/10.1016.j.brat.2021.103874.

Gawande, A. (2009). *The Checklist Manifesto: How to get things right*. New York: Henry Holt and Company.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischman, R. L., Hill, C. L., … Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale I: Development, use, and reliability. *Archives of General Psychiatry*, *46*, 1006–1011.

Grilo, C. M., White, M. A., Masheb, R. M., & Gueorguieva, R. (2015). Predicting meaningful outcomes to medication and self-help treatments for binge-eating disorder in primary care: The significance of early rapid response. *Journal of Consulting and Clinical Psychology*. https://doi.org/10.1037/a0038635

Jensen, A., Fee, C., Miles, A. L., Beckner, V. L., Owen, D., & Persons, J. B. (2020). Congruence of patient takeaways and homework assignment content predicts homework compliance in psychotherapy. *Behavior Therapy*, *51*(3), 424–433. https://doi.org/10.1016/j.beth.2019.07.005

Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, *63*(3), 146.

Kimball, R., & Caserta, J. (2004). *The Data Warehouse ETL Toolkit: Practical techniques for extracting, cleaning, conforming, and delivering data (1st ed.).* Wiley.

Kraus, D. R., Bentley, J. H., Alexander, P. C., Boswell, J. F., Constantino, M. J., Baxter, E. E., & Castonguay, L. G. (2016). Predicting therapist effectiveness from their own practice-based evidence. *Journal of Consulting and Clinical Psychology*, *84*(6), 473–483. https://doi.org/10.1037/ccp0000083

Kring, A. M., Persons, J. B., & Thomas, C. (2007). Changes in affect during treatment for depression and anxiety. *Behaviour Research and Therapy*, *45*, 1753–1764.

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ‑9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613.

LeJeune, J. T., & Luoma, J. B. (2015). The integrated scientist-practitioner: A new model for combining research and clinical practice in fee-for-service settings. *Professional Psychology: Research and Practice*, *46*(6), 421.

LeCompte, L. L., & Young, S. Y. (2020). Revised Common Rule changes to the consent process and consent form. *Ochsner Journal* 20(1): 62–75. doi: 10.31486/toj.19.0055

Lewis, C. C., Boyd, M., Puspitasari, A., Navarro, E., Howard, J., Kassab, H., … Douglas, S. (2018). Implementing measurement-based care in behavioral health: A review. *JAMA Psychiatry*.

Lovibond, S. H., & Lovibond, P. F. (1995). *Manual for the Depression Anxiety Stress Scales* (Second). Sydney: Psychology Foundation.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, *28*, 487–495.

Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The Work and Social Adjustment Scale: A simple measure of impairment in functioning. *The British Journal of Psychiatry*, *180*(5), 461–464. https://doi.org/10.1192/bjp.180.5.461

National Association of Social Workers. (2008). *NASW code of ethics*.

Obsessive Compulsive Cognitions Working Group. (2003). Psychometric validation of the obsessive beliefs questionnaire and the interpretation of intrusions inventory: Part I. *Behaviour Research and Therapy*, *41*(8), 863–878.

Osborne, T. L. (2018). Research in private practice settings. In *Practice-Based research: A guideline for clinicians* (1st ed., pp. 130–152). New York: Routledge.

Persons, J. B. (2007). Psychotherapists collect data during routine clinical work that can contribute to knowledge about mechanisms of change in psychotherapy. *Clinical Psychology: Science and Practice*, *14*(3), 244–246.

Persons, J. B., Bostrom, A., & Bertagnolli, A. (1999). Results of randomized controlled trials of cognitive therapy for depression generalize to private practice. *Cognitive Therapy and Research*, *23*, 535–548. https://doi.org/10.1023/A:1018724505659

Persons, J. B., Hong, J., Lemle Beckner, V., Owen, D., & Eidelman, P. (2012). *Monitoring therapy process at every session: Development of a new tool and clinical examples of its utility*. Association for Behavioral and Cognitive Therapies, Naval Harbor, MD.

Persons, J. B., & Thomas, C. (2019). Symptom severity at week 4 of Cognitive-Behavior Therapy predicts depression remission. *Behavior Therapy*, *4*, 791–802. 10.1016/j.beth.2018.12.002

Persons, J. B. (2021). *How to conduct research in your private practice*. Manuscript submitted for publication.

Persons, J. B., Osborne, T. L., & Codd, R. T. (2021). Ethical and legal guidance for mental health practitioners who wish to conduct research in a private practice setting. *Behavior Therapy*, *52*(2), 313–323. https://doi.org/10.1016/j.beth.2020.04.012

Persons, J. B., Roberts, N. A., Zalecki, C. A., & Brechwald, W. A. G. (2006). Naturalistic outcome of case formulation-driven cognitive-behavior therapy for anxious depressed outpatients. *Behaviour Research and Therapy*, *44*(7), 1041–1051. https://doi.org/10.1016/j.brat.2005.08.005

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097.

Vittengl, J. R., Clark, L. A., Thase, M. E., & Jarrett, R. B. (2019). Estimating outcome probabilities from early symptom changes in cognitive therapy for recurrent depression. *Journal of Consulting and Clinical Psychology*, *87*(6), 510.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Weissman, A. N., & Beck, A. T. (1978). *Development and validation of the Dysfunctional Attitude Scale*. Annual Meeting of the Association for Advancement of Behavior Therapy, Chicago.

Wilson, K. A., Persons, J. B., Reiser, R. P., Osborne, T. L., & Rizvi, Shireen L. (2019). *Keeping up to date with scientific advances: A practical guide for practitioners*. *the Behavior Therapist, 42(7),* 230-234.

Youngstrom, E. A., Choukas-Bradley, S., Calhoun, C. D., & Jensen-Doss, A. (2015). Clinical guide to the evidence-based assessment approach to diagnosis and treatment. *Cognitive and Behavioral Practice*, *22*(1), 20–35. https://doi.org/10.1016/j.cbpra.2013.12.005

Zieve, G. G., Persons, J. B., & Yu, L. A. D. (2019). The relationship between dropout and outcome in naturalistic cognitive behavior therapy. *Behavior Therapy*, *50*(1), 189–199. https://doi.org/10.1016/j.beth.2018.05.004

Appendix. Publications Based on the Naturalistic CBT Archival Database

Persons, J. B., Burns, D. D., & Perloff, J. M. (1988). Predictors of dropout and outcome in cognitive therapy for depression in a private practice setting. *Cognitive Therapy and Research*, *12*(6), 557–575.

Persons, J. B., Bostrom, A., & Bertagnolli, A. (1999). Results of randomized controlled trials of cognitive therapy for depression generalize to private practice. *Cognitive Therapy and Research, 23*, 535-548.

Persons, J. B., Roberts, N. A., & Zalecki, C. A. (2003). Anxiety and depression change together during treatment. *Behavior Therapy, 34,* 149-163.

Persons, J. B., Roberts, N. A., Zalecki, C. A., & Brechwald, W. A. (2006). Naturalistic outcome of case formulation-driven cognitive-behavior therapy for anxious depressed outpatients. *Behaviour Research and Therapy*, *44*(7), 1041–1051.

Kring, A. M., Persons, J. B., & Thomas, C. (2007). Changes in affect during treatment for depression and anxiety. *Behaviour Research and Therapy*, 45, 1753–1764.

Thomas, C., & Persons, J. B. (2013). Sudden gains can occur in psychotherapy even when the pattern of change is gradual. *Clinical Psychology: Science and Practice*, 20, 127-142.

Persons, J. B., & Thomas, C. (2019). Symptom severity at week four of cognitive behavior therapy predicts depression remission. *Behavior Therapy*, *50*, 791-802.

Zieve, G. G., Persons, J. B., & Yu, L. A. D. (2019). The relationship between dropout and outcome in naturalistic Cognitive Behavior Therapy. *Behavior Therapy, 50*(1), 189-199.

Gates, V., Hsiao, M., Zieve, G. G., Courry, R., & Persons, J. B. (2021) Relationship to CBT outcome and dropout of decision support tools of the written case formulation, list of treatment goals, and plot of symptom scores. *Behaviour Research and Therapy, 142.*

An additional study is underway:

Coe-Odess, S., Thomas, C., & Persons, J. B. (2020, November). *A longitudinal analysis of the performance of one evidence-based psychotherapist*. Presented at the virtual conference of the Association for Behavioral and Cognitive Therapies.