**KTDRR and Campbell Collaboration Research Evidence Training:**

**Management/Analysis Tools for Reviews - Abstrackr**

*Presenter: Ethan Balk*

A webcast aired on April 17, 2019, sponsored by AIR’s Center on

Knowledge Translation for Disability and Rehabilitation Research (KTDRR)

<https://ktdrr.org/training/webcasts/webcast67/index.html>

Transcript for audio/video file on YouTube:

<https://youtu.be/qzE9bqn8G-s>

JOANN STARKS: Hello and welcome to today's webcast brought to you by the Center on Knowledge Translation for Disability & Rehabilitation Research, or KTDRR, at American Institutes for Research, in coordination with the Campbell Collaboration. The Center on KTDRR is funded by the National Institute on Disability, Independent Living, and Rehabilitation Research, known as NIDILRR, in the US Department of Health and Human Services, Administration for Community Living. The Campbell Collaboration is an international organization that promotes positive change through the production and use of systematic reviews and other evidence synthesis for evidence‑based policy and practice. The Center on KTDRR partners with Campbell's Disability Coordinating Group to help increase the number of Campbell systematic reviews in the disability field.

I am Joann Starks with the Austin office of American Institutes for Research, or AIR, and I will be the moderator today. I want to thank my colleagues, Shoshana Rabinovsky and Ariana Hammersmith for helping with the logistics. The KTDRR Center and the Campbell Collaboration are working together to offer a 5‑part training course that focuses on high‑quality methods for synthesis of evidence, including the procedures and methods for conducting systematic reviews as well as software, tools and strategies for analyzing and reporting data.

Today we will learn about several software tools for managing and analyzing data in systematic reviews. We will hear about Abstrackr. Dr. Ethan Balk is an Associate Professor of Health Services, Policy & Practice at Brown University's School of Public Health. An internist by training, he has done extensive applied and methodological research in systematic review and meta‑analysis and he will start our presentation and talk about Abstrackr.

ETHAN BALK: I am going to be presenting Abstrackr citation screening software developed by members of Brown Center of Health, Brown University's School of Public Health. Our research for this tool was funded by the Agency for Healthcare Research Quality and you should feel free to contact me by email at any time regarding any questions about this. The link for the software is on the title slide here.

I also want to mention that none of the people working on this have financial conflicts of interest regarding this software beyond receiving research funds Abstrackr is an open access web‑based free platform used to screen publication citations or systematic reviews. It uses machine learning algorithms to predict acceptance of the unscreened abstract. I will talk more about that. The machine learning allows semi‑automation of the screening process, improving efficiency. There is the ability to annotate citations and also to organize the citations by acceptance status, by tags and by other factors.

Again, this was developed by our group at Brown Evidence‑Based Practice Center and Center for Evidence Synthesis in Health. We are a research group dedicated to conducting systematic reviews, improving the methodology around systematic reviews and related research. So the major features of Abstrackr: It shows screen presentation of the citations. It can be used easily on regular computer screens, on tablets, on smartphones. It allows simultaneous multiple screening. For example, a duplicate or screening duplicate or during pilot rounds when people are being trained about the endbility criteria and screening criteria. There is an ability to color‑code and rate the relevance of terms in titles, abstracts and key words, words and phrases.

There is the option to tag citation and to add notes about each of the citations. There is the software algorithm that allows prediction of potential relevance of the unscreened citations. This allows a number of different things, including sorting the unscreened abstracts in order of potential relevance. This allows front‑loading of the abstracts, the citations most likely to get accepted. So, relatively early in the process of abstract screening, one is likely to capture the large majority of abstracts that will end up getting accepted. We are still working on the empirical evidence to support and validate this but there is the potential to stop screening early or stop the full double screening of abstracts early based on the predictions and I will talk a little more about that so this is a screen shot of what Abstrackr looks like. With some annotations. I guess we lost some of the features here. What you can see is there is the full title, the journal information, author list. Full abstract and the list of key words.

If the abstract is long, then the screen, one would have to scroll up and down the screen. As you can see from the different colors of the text words, there is the ability to annotate using clock the title, abstract and key words. So, for example, this is a review about the substance use disorder treatments among children and adolescents.

So, we coded the study designs in purple, the random; the population can business and youth in green, the potential intervention in blue; and the reject ‑‑ some phrases that might indicate a higher likelihood that the abstract would be rejected in red: "Adult" in this case.

On the left there are buttons that allow one to tag the study, meaning to kind of give a short term that gets added to the citation and can be captured later on. And also the ability to add some notes. Down near the bottom of the screen are the Accept and Reject buttons and also a third button to allow one to say Maybe or defer a decision.

When setting up the project one has a number of different options. This is the screen that shows the process for entering a project into Abstrackr. Obviously, give it a name. But more importantly, there are options for either double screening or single screening; and there are options about how the abstracts will be sorted for presentation to the user. Most commonly, people would like to use the option to sort by relevance based on the prediction algorithm but there is also the possibility of presenting abstracts in random order or ignoring the algorithm relevance.

There is also the option to set up what are called pilot rounds. And these can be of any size. So in this setting, the way it's shown now, the team would ‑‑ all members of the team would see the same 100 abstracts; and regardless of how many people there are; and they would all screen those abstracts and the idea being that once that is completed, the team would get together and can go over the conflicts and maybe see who is an outlier, see what kinds of questions arise from that process and the idea is to get the team trained to all understand the eligibility criteria in the same way.

If a second round of screening is required, the number would be upped to 200 and everyone would do a second screening. After the pilot round, the screening would revert to the double or single screening options where everybody is presented with the citations. If it is set to double screening, after two people have been given a label for citation, no additional people would see that citation. The presentation of the double screening is done in the background. There is no assignment that needs to be made explicitly.

One needs to get the citations up into Abstrackr. There are a number of ways to upload the files or import the data into Abstrackr. The simplest is to develop a list of citations in EndNote or Reference Manager and then export out of EndNote as an RIS file and that can be directly imported into Abstrackr. There are some instructions on how to do that.

Alternatively, if one is using a list of citations that are coming exclusively from Pub Med, one can have a list of the Pub Med list of ID's there and populate that field. There is also a ‑‑ creating a tab‑delimited file for an Excel spreadsheet and uploading those files if desired.

I mentioned the concept of conflict resolution. The people who have screened in this case double‑screen abstract, if they disagree with each other, Abstrackr will note that, compile that list of abstracts together and then either a third person or the team as a whole can go back in and look at that list of conflicts.

In this example, there are 35 conflicts as indicated at bottom. You see that Gayle rejected and I accepted. The presentation of the abstract is nearly identical to the normal presentation of the citation. Really, just with the admission of how people screened the abstracts.

Once the team makes a resolution you can decide whether to accept or reject the abstract. I mentioned the prediction algorithm. So Abstrackr in the background uses machine learning algorithms to evaluate the abstract the that have not received any labels, have not been accepted or rejected. Based on the corpus of those that have been screened and the labels given to those, accept or reject ‑‑ actually it is a series of algorithms run to predict whether each of the unscreened abstracts are likely to be relevant.

Abstrackr then assigns a value between zero and one to each of those citations. This value is not a percentage or proportion likelihood that the abstract will be accepted; it's just a way to rank the abstracts. This is just one example from the same project fairly early on. You could see it's making a prediction as to how many citations are probably relevant. This is a very high prediction. But you can also how the remaining unseen abstracts are. The higher the number, the higher the prediction that the abstract will be accepted. In our experience, abstracts with a prediction value of less than 0.4 are extremely unlikely to be accepted.

As this distribution shifts to the left, the remaining abstracts are less and less likely getting accepted. Again, when one goes back into Abstrackr, starts screening again, one will be presented if this option shows, with the citations in this order. So, those abstracts most likely to be relevant get screened first. Just as a note for people trying to figure out system, this process of determining these prediction values happens once a night. It doesn't happen on the fly. It's actually a fairly time‑consuming, resource‑consuming process. This happens once a night based on everybody's predictions.

So, usually in the morning you get the abstracts sorted again and you see a rush of acceptance and then tapes off and then the next day kind of another rush of abstracts accepted that then tapers off. There are a number of ways of exporting. Probably the simplest and most common wave exporting out of Abstrackr is in the tab‑delimited file that can then be looked at in Excel or other spreadsheet database. We added in this color coded just for clarification. But what you can see is there is a study identifier information; any tags people have added. Any notes people have added for each of the citation. You could see everybody's label whether they rejected, negative 1. Accepted, positive 1. We are not shown an abstract that is a zero.

I guess I don't have an example. If somebody choose a 0 it would have shown up as 0. There is a column here showing there is convention consensus and what that consensus is. If there is some disagreement, then that shows up as an X and these would be the abstracts that get present the during the conflict resolution. This is not a complete database which is why the conflict here is not apparent.

So, some benefits of Abstrackr: We believe and we have some good experience that Abstrackr greatly increase the efficiency of abstract screening. What is very time‑consuming, sometimes the most time‑consuming step in a large review. The increases in efficiency are largely related to the ability of Abstrackr to predict likely relevance of unscreened abstracts and thus to front‑load the most likely to be relevant abstracts, which then allow one to capture those abstracts that are going to be accepted early in the process.

Again, as I mentioned, in our experience when all remaining prediction values are under 0.4, the screening process becomes extremely rapid because just looking at titles it becomes obvious that basically almost every if not every remaining abstract is going to be rejected, ineligible. We are still working on validating this, but we hope to be able to more conclusively say in what situations it's safe to stop screening early, which would be a great boon if, for example, in a corpus of 10,000 citations, one didn't need to screen the last 4‑5,000 citation. The tool, the Abstrackr can used with the machine learning features turned off if for some reason the group prefers that approach.

Some other functionality: If one does an update of exactly the same search you can add the new citation to the existing review. The advantage of that is that this would allow immediate predictions to occur. Normally, it takes several hundred ‑‑ it takes the screening of several hundred citations before the prediction algorithms start to kick in. But if you already have a full corpus with the completed prediction algorithm, then those algorithms can be applied to the new citations being added and one will immediately see the most likely to be relevant citation. And also, one can track the old and the new citation together in one file.

There are multiple ways of exporting out of Abstrackr. One that I didn't mention earlier is that one can now export directly into EndNote. So, you can export the labels back into EndNote, exports that have been accepted into a file at the labels. We are still working on the functionality of this a little bit. So, we should be able soon to be able to fully export the tags, the notes and so forth. There is also the option of exporting as an XML file for those groups that deem that to be useful.

We will soon be incorporating Abstrackr into another database tool SRDR+ repository that also is developed and managed here at the Brown Evidence‑Based Practice Center. SRDR is a tool primarily for test extraction and archiving extracted data. So, we are going to incorporate Abstrackr into this debase which will allow for full text screening within Abstrackr, development of evidence map or scoping maps and then final data extraction, all organized together.

Some groups require that a higher level of privacy of the data. So, it is possible ‑‑ it is an open‑source software; so, it is possible to load the software behind one's firewall, which would maintain the databases completely privately behind the firewall and we are happy to help groups with that. Some other improvements to come, as I mentioned, we will be transferring Abstrackr into SRDR+. I give the URL for that. This will streamline the various stages of abstract screening, full‑text screening, evidence mapping and eventually meta‑analysis and exporting tool. As I mentioned, there will be some further confirmation of validity of the prediction algorithms and the ability to use those prediction values to stop double or all screening.

JOANN STARKS: We have a question from the audience: That is: What kind of mapping output will be possible from SRDR+?

ETHAN BALK: We are starting up a section in SRDR. Plus, it should be available fairly soon where one can do full text screening of abstracts and then either as a separate processor at the same time answer a number of data questions, simple data extraction; for example, what's a study design, how many people in the study, what are the interventions, what are the outcomes looked at. The idea is to just ask simple, basic questions and then those data would be organized as something that could be evaluated within SRDR then exported out of it. Currently SRDR, the export or output process is again as a spreadsheet, Excel file. So, one can then manipulate the data in Excel.

In summary, Abstrackr is simple to use, simple to learn, intuitive, free platform with the primary purpose of screening citations. And it's usable pretty much on any web‑enabled device. It allows single or double training, pilot, conflict resolution. There is the ability to mark up, improve screening accuracy.

There are prediction algorithms available which again allows frontloading of the most relevant citations to more quickly find those eligible studies. It may allow early stopping of screening once all of the machine‑predicted accepted citation have been screened.

And updated literature searches can be incorporated seamlessly in Abstrackr to allow those to become very efficient processes. There is easy to use and multiple routes of importing and exporting data and importing the citation and exporting the labels of the citations and as I mentioned we will be incorporating in Abstrackr track are and SRDR for the screening, mapping and data extraction. I believe that is my last slide. We have included a number of references regarding the evaluation, validation of Abstrackr. Thank you very much. Here's some contact information.

JOANN STARKS: I want to thank our presenters for taking time to prepare and to introduce these software tools to help manage and analyze data for systematic reviews. We hope you will take a few minutes to give us some feedback about the webcast by filling out a brief evaluation—the link is listed in the slides. You only need to respond once, after you have viewed all 4 videos. I also want to thank the AIR staff and representatives from the Campbell Collaboration who helped with planning and logistics, and of course, we want to thank NIDILRR for their support to offer these webcasts and other events.