**CTSpedia Final Report:**
**September, 2012**

***Overview:***
This is the final report for a one-year administrative supplement to enhance the content and complete the administrative transfer of the CTSpedia platform to Vanderbilt University. We will review briefly the goals of the award and then summarize progress and accomplishments for each Specific Aim.

***Goals of the Administrative Supplement:***
CTSpedia (http://www.ctspedia.org) is an interdisciplinary collaborative resource for clinical and translational scientists and trainees, developed by a multi-site national CTSA-sponsored partnership. The goals of this supplemental awardwere ***to support the CTSpedia initiative activities*** and ***to transition CTSpedia to a sustainable resource for clinical and translational training and research.*** The award has directly supported activities of the Biostatistics, Ethics and Research Design (BERD) units at the four CTSA sites that have led the development of the CTSpedia's website and content: University of California, San Francisco (UCSF); University of Rochester; Vanderbilt University; and University of California, Davis. The primary award site for this initiative, following our rotation among CTSA sites, has been at UC Davis.

The supplement was awarded under Designated Topic 2, because CTSpedia is a key "legacy" resource developed by a multi-site CTSA partnership that required funding to facilitate its transition to a sustainable resource. We proposed in this transitional year to address three Specific Aims:

Aim 1. To complete and integrate the new content initiatives which have been added to CTSpedia portfolio since our ARRA funding but thatwere not covered in those administrative supplements. The initiatives included:

A. Expanded educational materials in residency training,
B. Statistical tools for clinical and translational research, and
C. Reproducible research tools and guidelines.

Aim 2. To enhance user-friendliness, accessibility, and accountability of CTSpedia for both content contributors and content users. These tasks included:

A. Improving site organization, links, and search functions,
B. Initiatives to increase our visibility and appeal to both potential users and potential contributors, in part through expanded content,
C. Evaluation both through one-time focused user surveys and ongoing tracking and feedback tools.

Aim 3. To transition to a sustainable, long-term model that does not require added supplemental funding. The two key tasks for the transition were:

A. Preparing the website infrastructure and administrative support at Vanderbilt for incorporation into a central coordinating center with minimal need for ongoing support,
B. Developing a template and guidelines for CTSA groups interested in developing new initiatives making use of the CTSpedia platform, to improve the likelihood that major new ideas can be successfully funded and implemented in the future.

***Progress report for Administrative Supplement (by Specific Aim)***

*Aim 1A: To develop and provide expanded educational materials for residency training.*
This component was carried out at UC Davis, by Dr. Beckett and a doctoral candidate in Biostatistics, supported as a graduate student researcher (GSR), with support from Dr. Banach.

1A.1. We developed and posted a modular set of units for "how to do a simple chart review project" designed to guide residents through the most common types of research projects using electronic or other medical record databases.
http://www.CTSPedia> >ResidencyProjectGuidelines
The resource outlines a for three-part sequence of activities, structured to be carried out over a standard three-year residency or, with more intensive work, in a shorter period. In Phase I the resident is guided in how to find a mentor, determine a question, develop a plan, obtain preliminary information for feasibility, make arrangements for using a data capture system such as REDCap, and prepare and submit an application to the IRB. In Phase II, the resident would collect general medical records data, determine data fields, organize the data, extract and code additional variables, and clean and document data for analysis. In Phase III, the resident would carry out statistical analysis and prepare manuscripts and presentations.

Drs. Beckett and Banach worked with GSR to identify suitable project examples, develop materials, and plan linkages to existing CTSpedia materials for each step of the three phases. We also identified IRB examples, data acquisition and data management tools and examples, and linked to the UCD CTSC informatics group for sample materials. Materials have been posted since spring and have been made available for testing and evaluation by UCD residents during their research training.

*Aim 1B: Develop statistical tools for clinical and translational research*
This component used contributions from Dr. Thurston at University of Rochester (ST) and from contributors at the FDA, as well as contributions from the other sites. Contributions and examples of their usage can be found both at the stat tools site and at the statistical graphics site within CTSpedia. More than 60 programs are currently available, from about 20 different contributors. There are also Stata Do files, instructions on how to use R scripts and SAS macros, and other useful general material.

Examples of specific functions contributed include:

- An R function to make a scatterplot, with superimposed adjusted or unadjusted regression lines (or both), with or without confidence intervals for the regression line. Example code in R and Sweave (the latter relates to reproducible research.

- An R function for graphics related to a regression with an interaction with one dummy variable (2 levels). The function plots y versus x and superimposes adjusted or unadjusted regression lines (or both) from the interaction. Example code in R and Sweave.

- A program "Mixed_msem": calculates mediation effect and 95% confidence interval for two level nested data structures, using mixed models.

- A program "Mixed_test_icc": calculates intraclass correlation and 95% confidence interval and tests if ICC is significant using mixed models.

- A program to calculate a U statistic based on ICC.
- A program fitting a proportional odds model.
- A SAS function: Zero-inflated Poisson model with > 2 assessments past baseline.

A second set of contributions has been provided by the Safety Graphics Group at the FDA, to share examples of typical graphics useful to the FDA in reviewing clinical questions. They have also written and linked a section on "How to Choose the Right Graph." Dale Plummer at Vanderbilt has helped Rich Forshee of the FDA group to implement his Stata graphics into GitHub gists. CTSpedia Safety Graphics will be featured in a new book on best practices in graphics for clinical research. This part of our Supplemental Award represents a unique collaboration between CTSC clinical and translational statisticians and FDA statisticians to share best practices to a wider audience, via the CTSpedia.

*Aim 1C. Reproducible research tools and guidelines.*

This project was carried out by the Reproducible Research Working group, under the direction of Dr. Harrell, with support from Dr. Banach.

The Reproducible Research Working Group outlined the contributions to be made to this section of CTSpedia. Among these topics, the group laid out a plan for contributions including: Standards and Definitions for Reproducible Research (RR), Papers on Reproducible Research, Reproducible Research Methodology, Reproducible Research Working Group Discussions and Mission Statement, and "What Went Wrong," reports on problems encountered in clinical research and how they did effect or could effect the study outcome. The working group then populated the RR portion of CTSpedia with technical material about RR implementation in statistical analysis and reporting. This important section has its own front-page section with links to web pages focused on each of the key topics above.

*Aim 2A. Improving site organization, links, and search functions*.
This aim was led by Dr. Banach, with able help from the technical staff at Vanderbilt, support from all four site leaders, and advice from the BERD Online Resources Working Group.

The success of this component is immediately clear from the new, user-friendly "front page" of CTSpedia at http://www.ctspedia.org The web page is physically attractive and well organized with high-profile links to search, jump, help, and log in. A sidebar has links to areas of the web for research ethics, the graphics working group, and a "sandbox" for experimentation. Drop-down menus connect to topics like education, reproducible research, and statistical tools and graphics. There are also links to suggestions for new material from contributors, and guides to help contributors get started. The search function is much more efficient, and the links are clear and easy to use.

*Aim 2B. Initiatives to increase our visibility and appeal to both potential users and potential contributors.*

All four site leaders have helped with this aim, with the leadership of Dr. Bacchetti, supported by advice and participation from the Online Resources Working Group. We have both expanded

the content and started new initiatives and programs to attract a wider group of users and contributors.

Dr. Bacchetti has expanded content in a number of directions. Among them:

- He has thoroughly edited and updated the CTSpedia article on common biostatistical problems to enhance its quality and usefulness.

- He has developed a checklist for what to look for in typical "Table 1" type output, summarizing characteristics of a study population. This is useful both to readers and to authors.

- He has audited and upgraded pages for hyperlinked column headings on enhanced macro output (e.g., logistic regression macro). Five entries have been added or updated. Another 10 desirable additions/modifications have been identified.

- He has identified new content areas to add, thus opening ideas for new contributors. These include developing an article or glossary page on estimation versus hypothesis testing and writing an article on logarithmic transformation, which has now been outlined and started (offline).

Two new initiatives have expanded the kind of discussion and resources available through the CTSpedia.

One initiative, led by Dr. Bacchetti and the BERD Online Resources working group with support from the BERD, was an online journal club. Two sessions have now been held, the first on "ethics in statistical practice" and the second on "p values and confidence intervals". Both had lively discussions, and these have now been posted and linked on CTSpedia. They have inspired 5 new topic discussion threads as follow-ups. This is a unique contribution to clinical and translational statisticians, made possible by the CTSpedia resource.

A second project of Dr. Bacchetti was to convert an article relevant to design of clinical and translational research, published as open access under the creative commons license, to a wiki article on CTSpedia. This has been completed. Conversion of another open access article by outside authors has been started.

An ongoing project at UCSF is to work with locally-funded programmers in UCSF's Biostatistical Consulting Unit to identify and adapt useful tools for posting at CTSpedia. Local programmers have completed an inventory of over 100 tools, utilities, and templates in use at UCSF. These require organization and conversion for use with CTSpedia subroutine calls instead of local calls. Dr. Bacchetti has piloted two initial projects for this, based on his logistic regression and on his Table 1 project, to ensure the same performance and understandability with CTSpedia's new form-based page structure for statistical tools as was present in the predecessor pages for these two tools. Local programmers at UCSF plan to follow with additional postings after this grant period.

All BERD members are invited to contribute BERD Success Stories to CTSpedia. These stories are then used for the CTSA Online Newsletter BERD Watch articles.

A brochure advertising CTSpedia was produced for the 2011 JSM Meeting. The brochure was quite popular and brought many new users to CTSpedia.

*Aim 2C. Evaluation both through one-time focused user surveys and ongoing tracking and feedback tools.*

Work on this aim has been carried out by Vanderbilt University CTSpedia staff in collaboration with Dr. Banach.

Dale Plummer has set up CTSpedia on Google Analytics. Monthly reports are downloaded to CTSpedia. Reporting is currently done for the most popular topic pages. Reports can also be done for specific topic pages. In addition, CTSpedia Foswiki software reports the highest number of topic page views and the contributors on a monthly basis. Recent Changes reports the topics that have been changed. Jeff Horner has re-programmed the Recent Changes to show the changes by timeframe. If one enters an individual's User Name in the Search, a report will be produced of all contributions a user has made on CTSpedia.

*Aim 3A. Preparing the website infrastructure and administrative support at Vanderbilt for incorporation into a central coordinating center with minimal need for ongoing support*.

This work has been carried out jointly by the Vanderbilt University CTSpedia group and Dr. Banach, with feedback from the other Supplement leaders and from other users.

Over the years of CTSpedia, consultants have added topic pages on how to use the wiki. All of the topic pages on "how to use the wiki" and contributing to the wiki have been collected in one spot with a primary location available on each page as well as tags to the relevant instructions.

Most contributions to CTSpedia, particularly tools, graphics, requesting services, templates, and references, are uploaded through what are called "forms," These provide an easily searchable database. Yinglin Xia, Rui Chen, and Sally Thurston from Rochester and Mat Soukup from the FDA have provided invaluable review of the forms and instructions for using the forms.

Jeff Horner uploaded R-Studio and GitHub to CTSpedia. R-Studio will help us with uploading and testing our scripts. Jeff uploaded Frank Harrell's Adverse Event Graphic to R-Studio and used a GitHub gist for version control for the script for Frank's Adverse Event Graphic. GitHub gists can be used for all of our attachments on CTSpedia and is a superior form of version control. This is critical for long-term stability and use.

*Aim 3B. Developing a template and guidelines for CTSA groups interested in developing new initiatives making use of the CTSpedia platform, to improve the likelihood that major new ideas can be successfully funded and implemented in the future.*

To make this aim successful, the leaders of the Supplement have worked closely with other CTSA working groups and their leaders, including those for the Online Research, the BERD Key Function, and others. Andy Cucchiera is working with the Clinical Research Management Group and Bob Oster is working with CTSA Education Group to see how we can help them and cross-link our groups.

With the help of Bridget Swindell at Vanderbilt we were able to video tape Jeff Hormer's training session on R-Studio and GitHub using GoToTraining. This format was highly successful and can be used by other groups wishing to post webinar materials on CTSpedia.

All working groups are asked to define the goals for their group. These are posted on CTSpedia.

Shelley Hurwitz will be chairing a new working group on Ethics in Clinical and Translation Science Biostatistics. The intent is to have webinars and contribute guidelines for ethical practice. The organization of this group will be used as the template for new working groups using CTSpedia.

***Summary:***

In summary, the Administrative Supplement team for CTSpedia has achieved the goals laid out in the proposal. The CTSpedia platform has successfully transitioned to Vanderbilt University, the organization and user-friendliness have been greatly enhanced, and the contents have been extended in many directions. In addition, we have initiated new projects that have expanded the user base and have opened doors for many future uses of CTSpedia. We hope that future support will be forthcoming to allow new projects and new partnerships, to follow up on the striking success of CTSpedia as a truly collaborative project in the spirit we believe NIH intended with the CTSA awards.