RESEARCH ARTICLE

A pragmatic trial of online interactive webtools for teaching biostatistics to first-year medical students: a constructivism-informed approach [version 1; peer review: awaiting peer review]

Steven Hardy, Douglas McHugh

Frank H. Netter MD School of Medicine, Quinnipiac University, North Haven, CT, 06473, USA

First published: 16 Aug 2021, 10:811
https://doi.org/10.12688/f1000research.55497.1

Abstract

Background: Statistical reasoning is an important clinical skill that informs evidence-based medicine and is desirable for 21st century practice. Yet many physicians struggle to apply biostatistical techniques correctly, potentially undermining patient care decision-making and adversely affecting outcomes. The overarching concepts of foundational biostatistics and probabilistic reasoning can be challenging to teach and hard for students to internalize given time constraints, lack of intrinsic motivation, and varied pre-existing knowledge or cognitive schema. We investigated how supplemental, interactive online webtools affected first-year medical students’ learning of sensitivity, specificity, negative/positive predictive values, and prevalence.

Methods: Online webtool design was informed by constructivism and sought to have learners 1) build conceptual representations by connecting new information to existing knowledge, 2) interact with primary sources of biostatistical data, and 3) begin with whole concepts then explore component parts. A total of 59 students were assigned randomly to two cohorts that had access to the webtools before or after in-class teaching; an additional 35 students had no access and served as the control cohort. Access metrics, summative exam scores, prior knowledge measurements, and qualitative data on learners’ experiences from three focus groups (n = 6/group) were collected.

Results: Mean exam scores were 87.5 ± 10.8%, 90.7 ± 11.2%, and 88.97 ± 12.9% for the before, after, and control cohorts (p = 0.49). Students repeatedly accessed the online webtools: mean access was 4.3 (± 2.6) (before) and 2.6 (± 1.5) (after). Qualitative data showed the webtools facilitated learning in a time-efficient manner.

Conclusions: Online interactive webtools facilitated a constructivism-informed learning model and were found by first-year medical
student to supplement and enhance their learning experience without detracting from established biostatistical teaching or negatively affecting summative assessment outcomes.

**Keywords**
Biostatistics, Medical Education, Shiny, R-software, Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value.
Introduction
The importance of understanding biostatistics in the context of communicating and health-related probabilities and applying them meaningfully with colleagues and patients has been highlighted as an important 21st century component of medical education. Indeed, doctors increasingly strive to practice evidence-based medicine and make clinical decisions that are supported by careful evaluation of peer-reviewed research. Therefore, an effective physician must be able to interpret biostatistical data competently when necessary. Unfortunately, practicing physicians struggle to parse sound knowledge from statistical data. To address this, many medical school curricula include dedicated biostatistics teaching grounded in medically relevant data sets. Some challenges that medical students face when pursuing proficiency in biostatistics include limited time, motivation, and prior knowledge or experience with which to scaffold learning. These challenges necessitate an engaging and streamlined curriculum so that students can get the most out of the effort they put in. Additionally, medical students tend to have a wide range of scheduling and study habits, so individualized learning is often preferred. These factors suggest that online resources are a good option to aide medical students in their pursuit of learning biostatistics due to the flexibility of asynchronous, digital instruction. One category of digital tools being increasingly used in teaching statistics is interactive graphics. The ability to adjust inputs and see immediate outputs (i.e., receive real-time feedback) enables fast perception of trends. This ability is especially useful in statistics when potential inputs can often range from single digits to numbers in the millions. Unfortunately, interactive graphics for statistical visualization are rarely applied to epidemiology or biostatistics in a medical school context. Instead, materials relying on problem sets and manual calculations predominate. Wider adoption of interactive graphics as a suitable teaching resource may be hastened by the generation of evidence to indicate that they are non-inferior to current educational methods. In this study, we investigated how supplemental, online interactive webtools affected first-year medical student learning of sensitivity, specificity, negative/positive predictive values, and prevalence – all of which are biostatistical concepts that concern the calculation and interpretation of probability information used in medical decision-making. We hypothesized that interactive online biostatistical webtools if offered as supplemental learning resources would be used by first-year medical students. We expected that the inherent properties of asynchronous, online resources (e.g., time-flexibility, fosters independent learning) would make the resources appealing to students. We also expected that a constructivism-based design would facilitate conceptual understanding of biostatistics content. We did not anticipate any decline in test scores of students using the new biostatistics resources.

Methods
Recruitment and consent
All 94 first-year students in the Frank Netter MD School of Medicine (FNHSOM) Class of 2023 were invited to participate in the study via an in-person announcement (given prior to a mandatory lecture), and then again via institutional email. Both the lecture announcement and email gave a brief overview of the study and indicated where students could find consent forms. If a student completed and submitted a consent form, they were considered enrolled in the study. The number of enrolled participants was 59. The Quinnipiac University IRB deemed data collection and analysis of those who did not enroll in this study, and therefore only interacted with the regular educational instructional strategies and practices for teaching and assessing competence with biostatistics at FNHSOM, to meet the exemption criteria for human subjects research (45 CFR 46.104(d)). These 35 students constituted the control cohort who had no access to the online interactive webtools.

Enrolled students were allocated into Before and After subgroups which determined whether they were able to access the online modules two weeks before or immediately after the regular in-person teaching event on ‘biostatistics used in medical decision-making’ in the Year 1 curriculum. Students at FNHSOM learn biostatistics in small groups of 5-6 students; membership of these small groups is determined through the random enroll feature of the Blackboard learning management system. Random assignment of Before or After status was made at this small group level for participants in order to ensure that they would mainly interact with others in their in-class biostatistics small group who had the same Before or After access status as themselves. For example, if four members of the six-person small group, “Biostatistics Small Group 1” consented to participate in the study, all 4 were assigned status Before; the remaining two students had no access to the interactive online tools. If three members of the six-person small group, “Biostatistics Small Group 2” consented to participate in the study, all 4 were assigned status After; the remaining three students had no access to the interactive online tools. Thus, three cohorts were constructed: the Before group of students had access to the online interactive webtools two weeks before in-person teaching by FNHSOM faculty occurred (n = 28); the After group of students had access to the online interactive tools immediately after in-person teaching by FNHSOM faculty occurred (n = 31); the No Access control group of students who only received in-person teaching by FNHSOM faculty (n = 35).

Focus groups were comprised of three sets of six students (18 total) who were selected randomly from the pool of study participants. Focus groups were 1 hour in duration and scheduled for the week after the FHSOM biostatistics summative exam so that students could reflect and report on the entire learning period of the study. Focus group questions were
centered around student motivation, tool usage, and differences in learning (e.g., “Given that using the biostats webtools was completely optional, what was your motivations to use them if you did or didn’t?”). All sessions were audio-recorded with participants using an assigned letter identifier (e.g., “Participant A”, and transcribed in a de-identified manner (by SH)).

Research approach
This study lasted for five weeks from August 15, 2019 – September 19, 2019 and was conducted through a mixed methods approach in order to combine “qualitative and quantitative viewpoints, data collection, analysis, and inference techniques for the purposes of breadth and depth of understanding and corroboration”. For the qualitative focus groups, we chose an interpretative phenomenological epistemology to explore participants’ experience and interpretation of using or not using the online interactive tools.

Intervention
The Before cohort were given access to the online interactive webtools two weeks before regular in-person teaching of ‘biostatistics used in medical decision-making’ at FHNSOM and for two weeks afterwards until the summative exam occurred. The After cohort were given access to the online interactive webtools immediately after regular in-person teaching of ‘biostatistics used in medical decision-making’ at FHNSOM and for two weeks afterwards until the summative exam occurred. The No Access cohort had no access to the online interactive tool at all but did participate in the regular in-person teaching of ‘biostatistics used in medical decision-making’ at FHNSOM and the summative exam.

Online interactive webtool development
Software
R software is a desktop application freely available from the Comprehensive R Archive Project (CRAN) that is well known and widely used in the statistical community. R statistical software was chosen as a platform because of SH’s pre-existing familiarity and its diverse statistical capability, online package availability, and open-source nature. A recent addition to its functionality has been the development of “Shiny” software that enables R to export a browser compatible HTML file whose functionality may be used to update the R code. By frequently resampling, the R software becomes “reactive” to its own HTML output, enabling an interactive interface. This means that a server with R software on it can be called to produce a webpage to any computer connected to the internet, and the webpage can be used to interact and manipulate with the server-based statistical coding (Figure 1).

Since its development, ShinyApps have been widely used, and have developed their own server hosting service enabling any member of the public to create an account, write code, and make it available for a large audience. Links to the ShinyApp webtool pages were displayed to students via the Quinnipiac University Blackboard® page. Blackboard® is a learning management system that allows online availability of electronic curricula as well as a host of analytic tools to enable tracking of web-link access.

Webpage creation
Using principles of instructional design and constructivism, a minimalist design approach was selected to create the interactive tool webtool pages. They had a simple title followed by simple graphics that displayed the biostatistical equations being explored (e.g., sensitivity) as well as a simple two-by-two table indicating relevant inputs. Graphics were screen captures from the regular FNHSOM biostatistics curriculum to signal continuity and cue students to the familiar content.

The second element of each webtool page was a set of instructions guiding students through the manipulation of input variables. These instructions directed attention to particular relationships. Questions were designed for students to quickly interact with the biostatistical equations in a way that encourages students to create predictions in a rapid estimate-based style and quickly check their results with the interactive tools for immediate feedback (Figure 2). Questions were written as plain text and placed in code that auto-formats based on the width of the users’ web-browser. An example of such a question is as follows

The third component of the webtool page was the interactive input bar, reactive two-by-two table, and refreshing calculated output values. Slider controls (Figure 3) were chosen for numerical inputs as they may add to the user’s concept of input magnitude. We created an additional feedback mechanism for the user by adding the movement and visualization of sliding numerical input because numerical literacy in the general population is a long-standing challenge and even those with substantial education experience difficulty interacting with, analyzing, and correctly interpreting...
straightforward numeracy questions. All slider inputs go directly into equations for the desired statistical parameter (i.e., $a/(a+b)$). Slider range was set to start at 1 to avoid confusing divide by zero situations, and preset to begin at the same value so that test values began at a neutral outcome.

The interactive two-by-two table (Figure 4) was created to show user inputs and aide understanding. Additionally, a `color_bar` function was added to represent the magnitude of inputs. Output values were simple reactive outputs `<renderText>` of inputs within the studied equation, rounded to 2 digits to minimize computational overload.

To promote learner progress from the ‘Understanding’ to ‘Applying’ and ‘Analyzing’ levels of Bloom’s cognitive taxonomy, the next portion of the webtool page was designed to allow translation of theoretical concepts to a scenario likely encountered by a first-year medical student (i.e., a peer-reviewed journal article from the PubMed® database, which contains 32 million searchable citations of biomedical literature). To do this, PubMed® was searched and article abstracts retrieved that presented relevant biostats concepts complete with reporting sample numbers and calculated results. These abstracts were presented as “Practice in Context” screen captures (Figure 5) within the webpage so as to be as realistic of an encounter as possible. A short set of questions pertaining to the displayed abstract tested the major concepts explored in the previous portion of the webtool page, with a similar style of estimate-based calculation. Some questions encouraged students to use the slider and reactive outputs as a simple calculator to check their work while tying applied scenarios back to basic concepts. Answers with a brief stepwise explanation were provided below for users to check their work and receive immediate feedback (Figure 6).
Data collection and analysis
The quantitative data collected consisted of online interactive webtool access metrics (i.e., the number of distinct visits to the relevant webpage by a participant), summative exam scores, and scores for a self-reported prior knowledge survey. An 11-question survey designed to query student prior knowledge of biostatistics used in medical decision-making was collected via Blackboard® before any in-person teaching or access to the online interactive webtools was granted (a copy of

\textbf{Figure 3. A screen capture of the input slider as seen in the sensitivity and specificity webtool.} Note: sliders are in their default position.

Data collection and analysis
The quantitative data collected consisted of online interactive webtool access metrics (i.e., the number of distinct visits to the relevant webpage by a participant), summative exam scores, and scores for a self-reported prior knowledge survey. An 11-question survey designed to query student prior knowledge of biostatistics used in medical decision-making was collected via Blackboard® before any in-person teaching or access to the online interactive webtools was granted (a copy
Figure 4. A screen capture of the reactive two-by-two table and corresponding positive and negative predictive value outputs as found in the positive predictive value and negative predictive value webtool. Note: the table is not showing default values to demonstrate the scaling color bars placed behind each input value.

|         | Disease | Healthy | Sum |
|---------|---------|---------|-----|
| Pos Test| 79      | 14      | 93  |
| Neg Test| 26      | 66      | 92  |
| Sum     | 105     | 80      |     |

Positive Predictive Value
0.85
Negative Predictive Value
0.72

Figure 5. A screen capture of the “practice in context” portion of the positive and negative predictive value webtool. Note: the complete screen capture of the abstract as displayed in PubMed®. Directly below the abstract image are the section instructions and a single practice question.

Using data pulled from the text above, consider the following questions, using the slider tools as a calculator for questions 3 and 4, but only when necessary in questions 1, 2, and 5:

Q1 In the results section, the authors report that the QuickRead go® produced 39 true positive results and 8 false positives. What is a simple ratio (like the ones mentioned above i.e. 1:2 1:3 etc.) that can be used to approximate the PPV calculation? How close is the approximation to the reported PPV?

Answers to Practice in Context
Q1 39:8 true pos: false pos -> 5:1 -> 5/(5+1) -> 0.83

Figure 6. A screen capture of answers provided for the “practice in context” portion of the Positive and negative predictive value webtool. Note: the brief notation showing each of the major steps in the cognitive process of coming to an answer denoted by arrows.
is provided as Extended data\textsuperscript{24}). Webtool usage data was collected from Blackboard\textsuperscript{®} through individual link clicks. Additional usage trends were observed from charts of server usage provided to the Shiny App administrator account used to host the online webtools. The summative assessment of students’ biostatistical knowledge occurred at the end of the month-long learning period. This exam consisted of 10 single-best answer multiple-choice questions, which were constructed to assess students’ competence of biostatistics used in medical decision-making, including definitions, calculation, and interpretation.

The quantitative data was imported to $s$ for analysis and descriptive statistics (e.g., mean, standard deviation) calculated. A two-sample independent Student’s t-test was used to compare the access metrics between the Before and After cohorts. One-Way ANOVA with Tukey’s HSD post hoc test was used to make pairwise comparisons of summative exam performance and prior knowledge scores between the three Before, After, and No Access cohorts. Analyses were conducted in SPSS v26 (IBM SPSS Statistics, Armonk, NY, USA) and the alpha level for statistical significance was set at 0.05.

SH facilitated the focus group sessions. De-identified audio recordings were transcribed manually then analyzed using an inductive, constant comparison approach.\textsuperscript{20} Major concepts were organized iteratively and stated as themes.

**Ethical considerations**
This study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Human Experimentation Committee/Institutional Review Board of Quinnipiac University (#08919; 22 August 2019). Written informed consent to participate was obtained from the participants, including consent to publish audio recordings of the focus groups.

**Results**

**Quantitative data**

**Prior knowledge survey**

The mean (± standard deviation) prior knowledge survey scores out of 110 points for the Before, After, and No Access cohorts were 53.8 (± 20.9), 53.6 (± 23.3), and 51.7 (± 22.1) points respectively (Figure 7). These scores were not significantly different from one another (p = 0.91; One-Way ANOVA) indicating that none of the three cohorts were different at the onset of the study. The range of the prior knowledge scores was 1 – 84 points. The median value was 55.5 points.

![Figure 7](image.png)

*Figure 7. A chart of the mean (± standard deviation) prior knowledge survey scores out of 110 points for the Before, After, and No Access cohorts. These scores were not significantly different from one another (p = 0.91; One-Way ANOVA).*
Summative exam scores

The mean (± standard deviation) summative exam scores as percentages for the Before, After, and No Access cohorts were 87.5% (± 10.8%), 90.7% (± 11.2%), and 88.9% (± 12.9%) respectively (Figure 8). These scores were not significantly different from one another (p = 0.49; One-Way ANOVA) indicating that none of the three cohorts were different at the onset of the study.

Figure 8. A chart of the mean (± standard deviation) summative exam scores as percentages for the Before, After, and No Access cohorts. These scores were not significantly different from one another (p = 0.49; One-Way ANOVA).

Access Metrics

Figure 9. A chart of the mean (± standard deviation) number of distinct visits per participant over the 5-week study period for the Before and After cohorts. These scores were significantly different from one another (p = 0.003; independent t-test).
Access metrics

The mean (± standard deviation) number of distinct visits per participant over the 5-week study period for the Before and After cohorts were 4.3 (± 2.6) and 2.6 (± 1.5) visits respectively (Figure 9). These scores were significantly different from one another (p = 0.003; independent t-test) indicating that the Before group chose to engage with the online interactive webtools more often than the After group. The range of the number of distinct visits per participant was 0 – 10 visits for the Before group, and 0 – 6 visits for the After group.

Server usage data available to the Shiny App administrator account showed a bimodal distribution over time, with the majority of student using the apps around the time during which the regular in-person ‘biostatistics used in medical decision-making’ was taught to the Class of 2023 (Figures 10 and 11). There was also a large spike in activity during the

![Sensitivity and Specificity App Use](image1)

**Figure 10.** A chart of the usage data for the sensitivity and specificity webtool for the 5 weeks over which students were taught, studied, and then assessed on their competence with biostatistics used in medical decision-making.

![Positive and Negative Predictive Value App Use](image2)

**Figure 11.** A chart of the usage data of the PPV and NPV webtool for the 5 weeks over which students were taught, studied, and then assessed on their competence with biostatistics used in medical decision-making.
few days before the summative exam. These major groups of activity were separated by approximately a week during which there was zero activity.

Qualitative data
Focus group discussion

Comparatively, about 75% of the focus group participants stated that they had opened the links provided to the online interactive webtools. In total, 50% said they used at least some portion of the webtools, while 25% reported engaging with and completing all interactive components of the webtools multiple times. A majority of students reported using the webtools for only one specific component. Some stated that they found the interactive calculators and tutorial most useful, while others found the practice-in-context element most useful. It was not uncommon for students to only use one and not the other.

Two themes from the focus group discussion emerged: facilitation of learning and time efficiency. Illustrative quotes are identified by the anonymous letter identifies assigned to each participant.

Theme 1 - Facilitation of learning

There was substantial agreement among the participants that the webtools enabled them to take more control of their learning processes to improve their comprehension and application of the biostatistical content.

A: “Once I hit a spot where I pause and I don’t understand, that’s when I used them [the webtools] for something extra to supplement my studying, something extra to help me understand.”

H: “I used the [webtools] as a way of tuning up for the exam. I wanted to make sure I was a bit more concrete and so it was nice to have something to walk through and play around with and get an actual feel for what that stuff looked like.”

C: “I was looking at the [regular curriculum teaching] slides and felt that they were a lot of definition and not numbers, so I used your [webtools] to play around with numbers and get an understanding of how sensitivity works.”

E: “I kinda struggled to understand some of the concepts or what we wanted to know about each concept from the [regular curriculum teaching] videos, so I used your [webtools] to play around and see how it would change without having to do each individual calculation.”

H: “What I thought was nice about the [webtools] is that you were able to … um … directly change the numbers and see what effect that had on the calculations without actually having to do all the calculations. So it saved a lot of leg work but you still got the big picture idea. Even though working out methodically can be helpful, it was nice to just see the trends.”

F: “I think I kinda use [the webtools] in conjunction with the [regular curriculum teaching] that we had and there are some examples with real studies and I kinda used that with the [the webtools] like ‘oh what would happen if I increased the number of people in this.”

Theme 2 - Time efficiency

There was also substantial consensus that the webtools helped the time investment needed to learn or set an appropriate learning pace.

D: “I like that I could have more control … like for me, the slow parts of traditional lectures are too slow for me, and the fast parts are too fast. [With the webtools] I have the ability to control the pace at which I go … allows me to be more effective and efficient.”

A: “I think that I would have been able to get to apply [my knowledge of the biostatistics] by spending a longer time with the lectures and rereading them, but I think that the [webtools] helped me get there faster than I would have. It was more efficient for me.”

D: “I would agree … that it kind of sped up the getting to applying and being able to broadly conceptualize it.”
D: “I would say that playing with the [webtools] did help me to get to apply faster.”

F: “I definitely would be more confident … just because I didn’t really understand the concepts behind the equations until I use [the webtools] so now I feel like I could explain this to my family if they were gonna get a test done that had those parameters.”

Discussion
This pragmatic trial study was designed to explore the use of interactive web-based resources to alleviate some of the challenges associated with teaching first-year medical students biostatistics such as a lack of time, motivation, and cognitive scaffolding. To accomplish this goal, constructivism was used as a conceptual framework to design interactive the webtools. Constructivism is a theory of learning that says, “people construct their own understanding and knowledge of the world, through experiencing things and reflecting on those experiences”.17 Put succinctly, it posits that learning is an active process whereby knowledge is constructed through appropriately contextualized experiences rather than mere transmission of information from teachers to learners. The webtools sought to have learners 1) build conceptual representations by connecting new information to existing knowledge, 2) interact with primary sources of biostatistical data, and 3) begin with whole concepts then explore component parts.21,22 The webtools were provided to the students for use alongside the regular biostatistics teaching curriculum at FHNSOM in two temporally spaced groups. Access metrics, summative exam scores, and qualitative focus group data was collected and analyzed. Our findings support our initial hypotheses that medical student users would interact with the webtools and deem them useful, that the constructivism-informed design would facilitate the conceptual understanding of learners, and that there would be no detrimental impact of using the webtools on biostatistical exam scores.

Quantitative lessons learned
There was no significant difference in summative exam scores or self-reported prior knowledge between the three study cohorts. These data, along with the interest expressed by the focus group participants and the repeated access metrics, indicate that this new resource could be added alongside existing biostatistics curricula without detrimental effects to learning. Results of access metrics data showed an elevated level of participation from those who chose to access the webtools with an average of 5 clicks per student. This indicated that many students found the webtool content valuable enough to return to after an initial experience. Furthermore, a temporal view of server data showed a significant spike of usage prior to the summative exam date suggesting that students returned to the webtools as a study resource during a high-pressure time period. Additional server data indicated that most webtool interactions were sessions less than 30 minutes in duration, suggesting that students were able to quickly visit the webtools to get what they needed rather than using them for long, unbroken sessions. This may be especially valuable to medical school students who are routinely trying to determine and make choices around what they will focus on learning in a time-scarce environment.

Qualitative lessons learned
Focus group discussion indicated a positive response to the constructivist design choices with students reporting that they allowed efficient use and increased students ability to experiment. Similarly, students endorsed the webtools’ part in allowing them to explain statistical concepts to others (e.g., friends or family members). Interestingly, one focus group argued against the constructivist design reporting that a more directed process would have been appreciated. This may reflect some learners’ preferences for passive versus active learning.23 The multi-component design and its associated flexibility of the online webtools was prominent in focus group discussion; a large portion of students revealed that they only used select portions of each webtool depending on their perceived personal learning deficits. Students often used the webtools as a way to test the knowledge and comprehension they had built and validate their conceptual understandings; our participants valued a setting where they may quickly experiment with concepts and receive immediate feedback – which is consistent with constructivism.

While we cannot claim that students had any increased learning motivation, it should be noted that once the webtools were created and made available to the Before and After cohorts, there was no additional reminder or external pressure to use them. No instructor time was required for their use; therefore, they may be cost effective resources regardless of how many students are motivated to use them. There was consensus in focus groups that given the low resources necessary to keep the webtools running they should be made available to future students to broaden their learning options.

Access data indicated that students mostly used the webtools during the weeks concurrent with the regular in-class curriculum at FHNSOM. This suggests that students may have desired additional or different types of resources during their active knowledge construction processes. This bore out in focus group discussions when participants mentioned regularly looking for a multitude of resources to be used at their discretion during periods of intensive academic focus.
In focus groups, some Before cohort students mentioned accessing the webtools to “pre-study” the material. They generally reported finding the webtools alone without introduction or explanation (provided by the regular biostatistics curriculum) confusing and counterproductive. Further, students who found the modules confusing reported abandoning them altogether or stopped using the webtools until a later date.

**Study limitations**
The limitations of this study include that it was conducted once with a small sample of learners at a single institution.

**Conclusion**
Most curricula for biostatistics follow some variation of the following steps: 1) knowledge acquisition via text or lecture format, 2) knowledge application via problem sets or other use-case based tasks, and 3) testing to assess learner competency. This sequence moves students along Bloom’s cognitive taxonomy from remembering > understanding > application > analysis.\(^{18}\) Due to the constraints of synchronous teaching modalities like timing of lectures or physical presence in student discussion groups, traditional methods offer limited opportunities for the learner to move up and down Bloom’s taxonomy as they might see fit and at their own pace. Through the use of R software and Shiny we created webtools for producing interactive, web-ready data visualizations and practicing learning in context. These resources were valued by learners for facilitating student-paced learning in a time efficient manner and may represent a pragmatic approach to supplementing existing medical school teaching strategies for biostatistics and learners’ abilities to curate their own learning experience.

**Data availability**

**Underlying data**
Zenodo: A Pragmatic Trial of Interactive Online Statistical Webtools for Teaching Biostatistics to First Year Medical Students: A Constructivism-Informed Approach. [https://doi.org/10.5281/zenodo.5092290](https://doi.org/10.5281/zenodo.5092290)

This project contains the following underlying data:

- Steven Hardy F1000 Data (Quantitative data).xlsx (access metrics, summative exam scores, prior knowledge survey scores).
- Focus Group 1 09-30 Participants A-E.m4a (focus group 1 audio recording).
- Focus Group 2 09-30 Participants F-H.m4a (focus group 2 audio recording).
- Focus Group 3 10-01 Participants I-N.m4a (focus group 3 audio recording).
- Focus Group 4 10-02 Participants O-R.m4a (focus group 4 audio recording).
- Focus Group 1 participants A-E.docx (focus group 1 transcript).
- Focus Group 2 F-H.docx (focus group 2 transcript).
- Focus Group 3 I-N.docx (focus group 3 transcript).
- Focus Group 4 O-R.docx (focus group 4 transcript).

**Extended data**
Zenodo: A Pragmatic Trial of Interactive Online Statistical Webtools for Teaching Biostatistics to First Year Medical Students: A Constructivism-Informed Approach. [https://doi.org/10.5281/zenodo.5092290](https://doi.org/10.5281/zenodo.5092290)

This project contains the following extended data:

- Prior Knowledge of Biostatistics Used in Medical Decision.docx (a blank copy of the self-reported ‘prior knowledge of biostatistics used in medical decision-making’ survey)

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).
References

1. Wartman SA: The Empirical Challenge of 21st-Century Medical Education. Acad Med. 2019; 94: 1412-1415. PubMed Abstract | Publisher Full Text
2. Rosenberg W, Donald A: Evidence based medicine: An Approach to clinical problem-solving. BMJ. 1995; 310: 1122. PubMed Abstract | Publisher Full Text | Free Full Text
3. Masic I, Miokovic M, Muhamedagic B: Evidence Based Medicine - New Approaches and Challenges. Acta Inform Med. 2008; 16: 219. PubMed Abstract | Publisher Full Text | Free Full Text
4. Gigerenzer G, Gray JAM, John AM, et al.: Better Doctors, Better Patients, Better Decisions: Envisioning Health Care. The MIT Press; 2011.
5. Wegwarth O: Statistical illiteracy in residents: what they do not learn today will hurt their patients tomorrow. J Grad Med Educ. 2013; 5: 340-341. PubMed Abstract | Publisher Full Text | Free Full Text
6. Astin J, Jenkins T, Moore L: Medical students’ perspective on the teaching of medical statistics in the undergraduate medical curriculum. Stat Med. 2002; 21(7): 1003-1006; discussion 1007. PubMed Abstract | Publisher Full Text
7. Fielding S, Poobalan A, Prescott G, et al.: Views of medical students: what, when and how do they want statistics taught? Scott Med J. 2015; 60(4): 164-169. Publisher Full Text
8. Seeing Theory: a visual introduction to probability and statistics. Accessed March 17, 2019. Reference Source
9. ArtofStat: Web Apps. Accessed July 26, 2021. Reference Source
10. Johnson BR, Onwuegbuzie AJ, Turner LA: Toward a definition of mixed methods research. J Mixed Methods Res. 2007; 1: 112-133. Publisher Full Text
11. Neubauer BE, Winkop CT, Varpio L: How phenomenology can help us learn from the experiences of others. Perspect Med Educ. 2019; 8: 90-97. PubMed Abstract | Publisher Full Text | Free Full Text
12. CRAN: The Comprehensive R Archive Network. Accessed July 26, 2021. Reference Source
13. R: The R Project for Statistical Computing. Accessed July 26, 2021. Reference Source
14. Shinyapps.io. Accessed July 26, 2021. Reference Source
15. Ettema PA, Newby TJ: Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. Performance Improvement Quarterly. 1993; 6: 50-72. Publisher Full Text
16. Paulus JA: Innumeracy: Mathematical Illiteracy and Its Consequences. New York, NY, USA: HOLT MCDOUGAL; 2001.
17. Lipkus IM, Samsa G, Rimer BK: General performance on a numeracy scale among highly ducated samples. Med Decis Making. 2001; 2: 37-44. PubMed Abstract | Publisher Full Text | Free Full Text
18. Anderson LW, Krathwohl DR, Airasian PW, et al.: A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. New York, NY: USA: Pearson; 2000. Reference Source
19. Butler-Kisber L: Constant comparison inquiry. In: Qualitative Inquiry. Butler-Kisber L, Ed.; London, UK: SAGE Publications Ltd; 2018; pp. 41–59.
20. Colburn A: Constructivism: Science Education's "Grand Unifying Theory". The Clearing House: A Journal of Educational Strategies, Issues and Ideas. 2000; 74: 9–12. Publisher Full Text
21. Fosnot CT: Constructivism: Theory, Perspectives, and Practice. New York, NY: USA: Teachers College Press; 2013.
22. Eva KW, Regehr G: "I'll never play professional football" and other fallacies of self-assessment. J Continuing Education Health Professions. 2008; 28: 14–19. Publisher Full Text
23. Hardy S, McHugh D: A Pragmatic Trial of Interactive Online Statistical Webtools for Teaching Biostatistics to First Year Medical Students: A Constructivism-Informed Approach [Data set]. Zenodo. 2021. Publisher Full Text
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com