

Visualizing Uncertainty: Helping Atrial Fibrillation Patients Understand That Risk Estimates Are Imprecise

James Hicklin

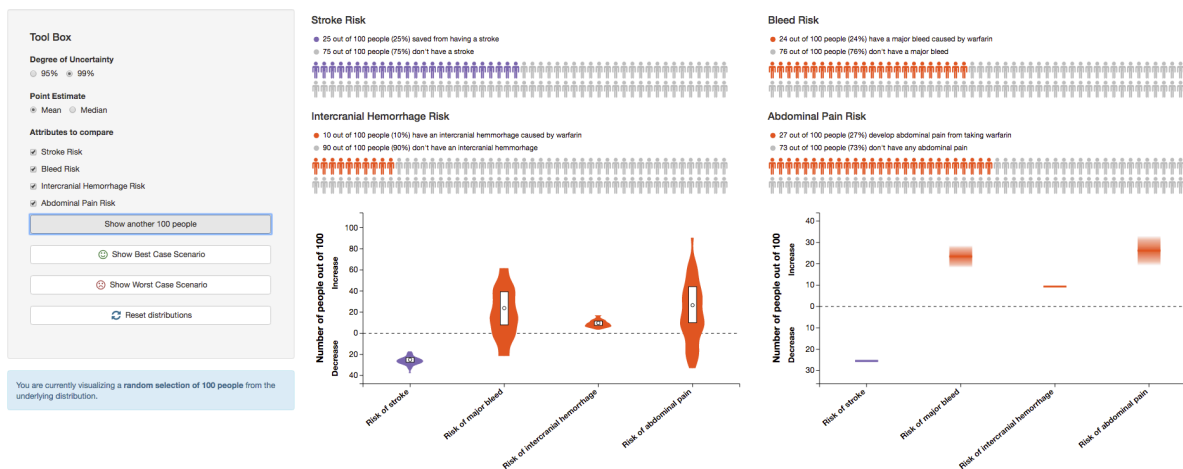


Fig. 1. Default view of the developed tool to understand uncertainty. The toolbox view appears on the left, while the visualization comparison view appears on the right.

Abstract—All risk estimates contain some degree of uncertainty, but this information is rarely, if ever, provided to patients. We present a novel tool with which developers of decision support tools can use to decide which visualizations best fit within their given disease context. We developed the tool in the context of atrial fibrillation, but the techniques used are generic in the sense that they could be used in other disease processes. We use violin plots, gradient charts, and icon arrays as representations of risk, and offer a toolbox that allows users to interact and adjust the parameters associated with the uncertainty around each risk estimate. The tool was deployed on our personal website on a free subdomain, and uses standard web technologies so that users can access the tool in a browser as opposed to downloading a standalone application.

1 INTRODUCTION

In recent years, the physician-patient relationship has dramatically changed, in that patients are encouraged to participate in therapy decisions related to their health issues^[1]. This phenomenon, termed shared decision-making, often follows a familiar formula:

1. Provide patients with some background information relating to their health issue.
2. Present patients with the available treatment options, including risk estimates related to the benefits and harms of treatment.
3. Ask patients to choose which treatment option they think best aligns with their personal preferences.

Often, this formula is implemented in the form of a decision support tool (DST), usually available online as an interactive online instrument. The process of weighing the pros and cons of potential treatment options inherently requires a cognitive trade-off, in which the potential benefits of treatment must outweigh the potential risks. While the International Patient Decision Aids Standards suggest including risk estimates in all DSTs^[2], there is always a degree of uncertainty when it comes to health, and it is crucial that this uncertainty be conveyed to the patient. Much like in the world of

finance, heterogeneity amongst the population produces individuals on the spectrum of risk-seeking to risk-averse, and having a thorough understanding of treatment-related risks would likely improve patient satisfaction with respect to their final treatment decision.

Conceptually, uncertainty can be difficult to understand without some form of visual aid. Typically, the literature differentiates first-order, or aleatory uncertainty to second-order, or epistemic uncertainty. First-order uncertainty is a form that arises due to an inherent variability in the general population. If we know that 5 out of 100 people will experience some side-effect, we do not know which 5 out of every 100 people it will be. This is an example of first-order uncertainty. Second-order uncertainty arises due to a lack of information regarding the point estimate. If we have a point estimate that indicates 5 out of every 100 people will experience some side-effect, it might actually be 3 or 7 out of every 100 people. In reality, point estimates will always have some confidence interval surrounding them, which corresponds to the second-order uncertainty in the risk estimate. In patient-oriented DSTs, the uncertainty surrounding a risk estimate is often provided as a simple textual representation, using phrases such as “about 5 in 100 people experience some side-effect”^[3]. While it is commendable that the developers of these tools acknowledge that there is uncertainty in the risk estimate, simple textual depictions are insufficient to properly convey the degree to which uncertainty affects the point estimate presented.

James Hicklin is with the University of British Columbia.

E-mail: hicklin.james@gmail.com.

Tool available at: <http://understandinguncertainty.jameshicklin.com>

Source code available at: https://github.com/hicklin-james/547_project

2 RELATED WORK

While the issue of uncertainty is often raised in visualization research, encoding it within visualizations has proven to be difficult. Research has identified that a major constraint to representing uncertainty is the additional dimension required^[4]. In the case of a single point estimate representing a risk, an error bar that conveys uncertainty around the y-value changes the data mark from a point to a line. Similarly, an isoline on a contour map would require an area mark to convey the uncertainty around it. As our proposed visualization is primarily intended as a tool to convey risk estimates, we limited the scope of our literature review to only examining how risk and risk uncertainty is visualized.

Johnson and Slovic^[5] were one of the first groups to investigate layperson understanding of uncertainty in risk estimates. While their research was done in the environmental risk domain as opposed to risk within a disease or health process, their findings were critical as they discovered that trust in risk estimates is significantly affected by the presentation of confidence intervals, and especially where those confidence intervals fall. Notably, their results indicated that ranges that included zero risk were trusted less than those that didn't. Additionally, participants did not appear to distribute equal probabilities of risk within the range of a confidence interval. They tended to believe values of higher risk within a range were more likely than those of lower risk. This lends evidence towards ambiguity aversion, in that individuals presented with ranges or confidence intervals focus on the extremes of those ranges. Additionally, this might indicate that encodings that properly represent the probability at specific points within the confidence interval might aid participants in understanding what the underlying distribution could look like.

One study by Han et al.^[6], looking at the effects of communicating uncertainty in individualized colorectal cancer risks, compared textual representations of uncertainty with gradient bars. Results indicated that when presented with gradient bars as opposed to simple point estimates, participants perceived their actual risk of cancer to be lower, which was contrary to what the authors hypothesized and inconsistent with the theory of ambiguity aversion; a phenomenon in which people avoid ambiguous or uncertain outcomes. This suggests that optimal representations of uncertainty may be able to better inform patients with respect to the underlying data, while simultaneously minimizing the issue of ambiguity aversion that has been well documented in the literature.

Correll and Gleicher^[7] investigated the drawbacks of using traditional error bars to convey uncertainty, and assessed how alternate encodings might improve understanding within a general audience. They hypothesized that gradient bars and violin plots, which better represent the probability distribution around a point estimate compared to confidence intervals, might lead to better inferences regarding the mean and error of the data. The results of their study indicated that encodings that portray more detailed information about the unlikely outcomes outside the margin of error improved the general audience's performance with respect to aligning their expectations with statistical expectations. Additionally, they found that encodings that displayed uncertainty in a continuous manner were more effective than those that displayed uncertainty as individual points, such as in confidence intervals. Overall, they found that the general population has the ability to make nuanced inferences from graphical data, even in cases where visual encodings, such as violin plots, are being experienced for the first time. This conclusion is vital, as it provides an encouraging baseline with which researchers can expand upon to implement new, and hopefully better visualizations.

An obvious solution to more generally visualize uncertainty is to incrementally add and encode the additional facets of uncertainty. Kao et al.^[8] do this in 3 dimensions, by deforming the 3D surface based on its standard deviation, encoding the interquartile range by adjusting color hue, and adding an additional line channel to encode the difference between mean and median. While this is an effective

approach in the expert use case, our solution needs to be simple and intuitive so that it is understandable from the perspective of a more general population, as the intended use of the visualizations are within a patient population. We therefore restrict our solution to 2 dimensions, and make no attempts to visually encode every aspect of uncertainty on the same plot.

Hakone et al.^[9] published a design study in which a patient-centred visualization for prostate cancer risk was designed and assessed. Their results indicated that visualization was an effective method to convey risk estimates in patients with low numeracy and visual literacy. They emphasized the importance of balancing readability and ease of comprehension, and warned against complex visualizations. Their results found that some patients were confused when using basic interaction elements such as sliders, and very few participants understood a relatively simple temporal area chart. This highlights the need for simple visualizations when they are to be presented to patients, especially in older population such as cancer and atrial fibrillation patients.

Bansback et al.^[3] conducted a review of current decision aids to determine how uncertainty is currently being presented. Their results indicated that 47% of decision aids reported no second-order uncertainty, while the rest reported at least a simple textual representation. This was significant, as it indicated that a large portion of DSTs reported risk estimates with implications of precision, when in fact, said estimates are always imprecise. While the degree of imprecision might vary, it is always worthwhile to report it as it may influence patient decision-making.

There exists some research that contradicts the general consensus that visualization improves user understanding of risk. Micallef et al.^[10], in their crowdsourcing study, found that simply adding a visualization to a textual Bayesian problem fails to improve user understanding of their individual risk, even when the text refers to the visualization. They did find, however, that visualizations can be helpful if the text contains no numerical values. They tested Euler diagrams, frequency grids (icon arrays), and hybrids of the two in their design, and found results that were inconsistent with existing literature. They emphasized the need for more research in this domain.

3 DOMAIN, TASKS, AND DATASET

3.1 Domain

Our domain of focus is in the context of someone developing a DST for atrial fibrillation patients. While the tool itself is intended for DST developers, the overarching goal is that the visualizations within the tool are eventually injected into the DST. As a result, it is vital that the visualizations themselves stay simple to a degree at which they can be understood by patients, while any interactivity and parameter adjustments made within the tool can assume some degree of familiarity with statistics and uncertainty.

In the patient context, significant trade-off must occur between the pros and cons of therapy in order for the decision-making to be optimal. In atrial fibrillation, patients must decide whether the potential benefits of warfarin in reducing stroke risk outweigh the risk of major side-effects, including internal bleeding and intracranial haemorrhage. The estimates around these risks and benefits always have some degree of uncertainty around them, yet this is rarely presented to the patient in an objective, easy-to-understand way. Yet, if the patient were to have access to this information, especially when the uncertainty is significant as is often the case in new therapies, the effect on the patient's decision-making may be enormous.

In the developer context, careful consideration must be taken when choosing which visualization to use in a DST. Visual literacy, numeracy, and technological competency must be accounted for, as isotopes may be suitable for one population while violin plots may be suitable for another. In the case of AF, the majority of patients are over 60 years old, so familiarity with uncommon plots such as

violin plots may dissuade the developer from using such encodings to present risk estimates.

3.2 Data

The input data for the tool is entirely synthetic, for reasons explained in the limitations. It consists of a JSON object, where each key at the root level corresponds to a specific attribute related to atrial fibrillation. Within each key lies an array of point estimates, which represents the percentage change in number of patients who experience the given risk or benefit. Note that data where the point estimate represents a benefit is negated, so that the direction of the data remains consistent, as the data is conceptually diverging. The data was synthesized using NumPy, as the library provides straightforward functions to sample from a given distribution. In the final tool, there were ~4500 unique point estimates across all the attributes, though this number can easily change by modifying and running the sampling file. The following shows a sample of the input data structure:

```
strokeRisk:
  data: [-24.5, -23.6, -25.4, ...]
bleedRisk:
  data: [45.5, 25.3, 34.7, ...]
```

While the input data is relatively simple, there was a significant amount of derivation required to compute the various variables associated with uncertainty. For each disease attribute, we calculated the minimum value, maximum value, inter quartile range, median, mean, 95% confidence interval, 99% confidence interval, and standard deviation. Only a subset of the derived data was used by individual visualizations within the tool.

3.3 Tasks

Given the domain identified previously, the primary goal of the solution is as a tool for developers of DSTs to compare various methods of visualizing uncertainty in risk estimates, and choose which encodings best fit within their decision support tools. Specific tasks include:

1. Compare violin plots, gradient plots, and isotype plots when visualizing the entire distribution. As each plot encodes uncertainty differently, this should help inform the developer as to which plots may be unsuitable.
2. Adjust the confidence interval of the presented uncertainty between 95% and 99%.
3. Adjust the point estimate presented, between mean and median values.
4. Filter attributes related to the treatment that are not relevant, to simulate how a patient may interact with the visualizations.
5. Randomly sample the underlying distribution for each risk estimates, to see how 100 actual patients may experience the risks and benefits.

Additionally, as each of the attributes governing the amount of uncertainty change, the visualizations should update in real time, so that developers can see what effect their interaction has on the visualizations.

4 SOLUTION

Our solution is intended to be used by developers of decision support tools who wish to include information regarding the risks and benefits of treatment options.

Based off of the tasks listed in the previous section, our design is split into two major components:

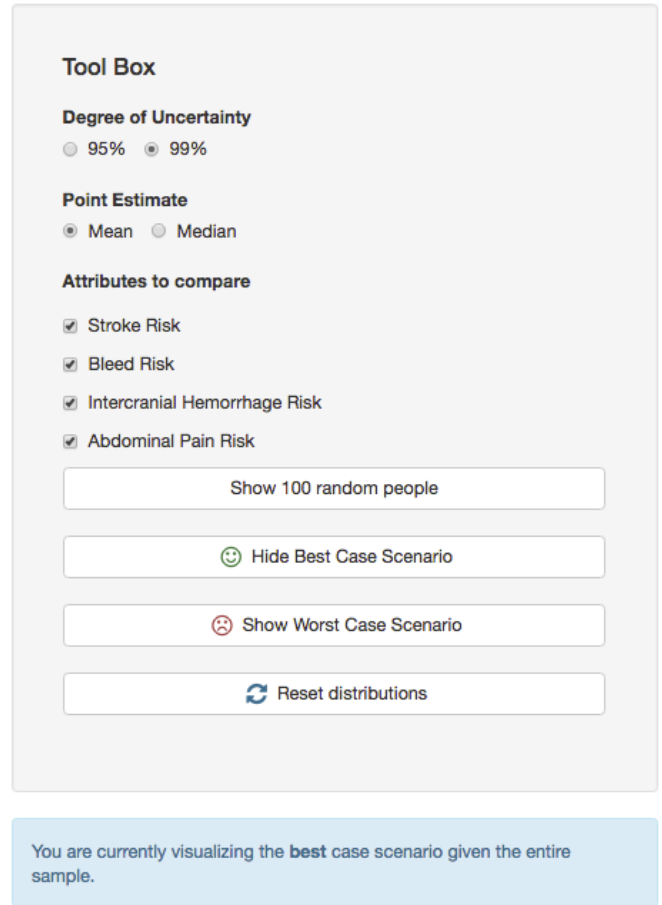


Fig 2. The toolbox view

- Toolbox view: This view contains all of the toggles and interactivity needed to adjust the parameters of uncertainty to be visualized.
- Comparison view: This view contains the visualizations of uncertainty within the same viewport, so that users can directly compare how modifications to the uncertainty affect the visualizations.

4.1 Toolbox view

The toolbox view appears on the leftmost side of the screen, and contains a variety of checkboxes, radio buttons, and standard buttons that allow users to interact with and modify the comparison view. A screenshot of this view is seen in Figure 2.

The toolbox view is the primary view in which interaction with tool happens. The following interaction is supported:

1. **Degree of uncertainty radio buttons**
This is where the confidence interval of the data is changed between 95% and 99%. This has a direct effect on the gradient charts, as the confidence interval grows or shrinks, and an indirect effect on the isotypes, as the best and worst case scenarios depend on the level of confidence.
2. **Point estimate radio buttons**
This is where the point estimate of the data can be changed between the mean and median. This has a direct effect on all three charts. The y position of the point estimate on both the violin plots and gradient plots

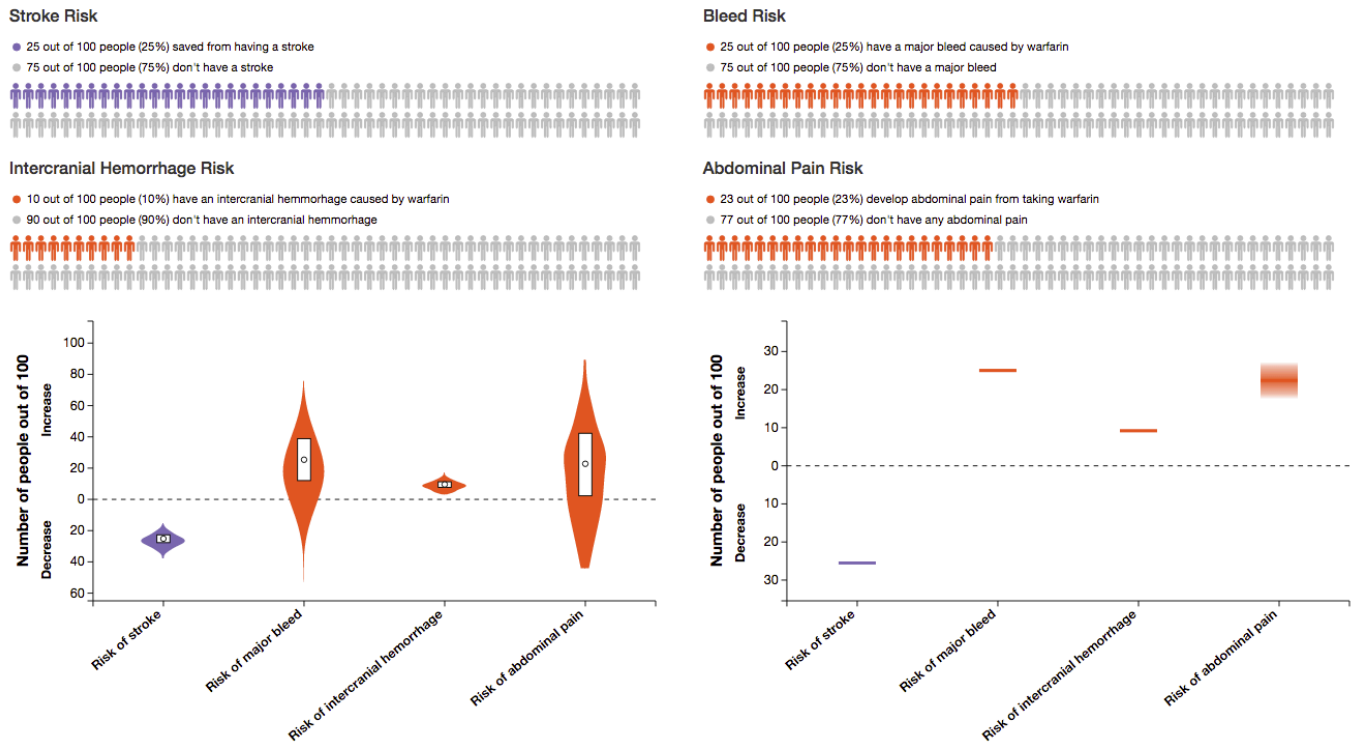


Fig 3. The comparison view. Isotypes, also called icon arrays, are stacked in groups of two at the top of the view. The violin plot is below the icon arrays on the left hand side of the screen, while the gradient charts are on the right hand side of the screen.

changes, while the number of people experiencing an effect on the isotype visualization changes.

3. Attributes to compare checkboxes

These buttons allow filtering of the data. In the default case, all attributes are presented across all visualizations. These can be filtered down by unchecking these boxes. All three charts will update dynamically as these are checked and unchecked.

4. Show 100 random people button

This button randomly selects 100 individual point estimates from the input data, and re-computes the statistical measures needed to draw the charts. Each visualization dynamically updates when this button is clicked.

5. Show best case scenario button

This button shows the best-case scenario given the currently set confidence interval. “Best-case”, in the tool, is the lower end of the current confidence interval.

6. Show worst case scenario button

This button shows the worst-case scenario given the currently set confidence interval. “Worst-case”, in the tool, is the upper end of the current confidence interval.

7. Reset distributions button

This button resets the charts so that the uncertainty data is computed based off of the entire sample.

An additional component of the toolbox view is a bootstrap alert underneath the toggles, informing the user of what their current selection of parameters represents. This is necessary to ensure that the user doesn't lose track of what filters and parameters they have applied up to this point in time.

4.2 Comparison view

The comparison view is where the visualizations themselves lie, and where users can compare how uncertainty is presented across

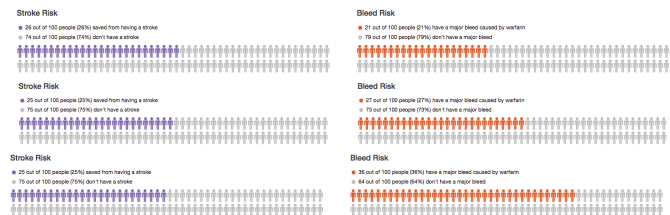


Fig 4. Three different 100-person samples from the same overall point estimate across two different attributes.

visualizations. They lie within the same viewport such that scrolling is unnecessary. Despite this, they are mobile friendly, and on smaller screens, will stack vertically. The entire comparison view can be seen in Figure 3.

Any time a user changes a parameter in the toolbox view, some, or all of the charts within the comparison view will update automatically to represent the new parameters. Color hue is used consistently across visualizations to encode the direction of the delta, and was selected based off of suggestions for encoding diverging data in Munzner's Visualization Analysis and Design [11].

4.2.1 Isotype visualization

Isotypes, also known as icon arrays in health literature, were selected as a primary visualization due to their reported success at conveying risk estimates in the health field, especially to older visually illiterate populations [12]. Reportedly, their success is due to their representation of risk being in the form of natural frequencies as opposed to percentage changes. Since a percentage change relies on understanding the percentage baseline, and then stacking the percentage delta on top of that baseline, it can be easily misunderstood, especially in non-experts of the general

population. Many DSTs include icon arrays as part of their presentation of risk estimates. In the case of traditional icon arrays, the point estimate is encoded by the number isotype marks, and the direction of the change is encoded by the color hue. Unfortunately, in their traditional static form, icon arrays offer no additional channel with which we can encode additional attributes, namely those associated with uncertainty. In our tool, we resolve this issue through the use of interaction.

Given some pair of estimates, where the point estimate is similar, representation with a static icon array may result in an incorrect conclusion that the point estimates are equally precise. Using the “Random 100 person sample” button in our tool, we can get a better idea of what that underlying distribution looks like.

In Figure 4, two attributes are being compared. The absolute value of the point estimate is equal between attributes, when comparing the overall sample. In other words, for every 100 patients, the number of people who are saved from having a stroke as a result of treatment is equal to the number of people who have a major bleed as a result of treatment. However, this point estimate is based off of the entire sample. If we randomly select 100 people from the underlying sample, and display those selected people on the icon array, we gain some insight into the underlying probability distribution. In Figure 4, despite the mean point estimate being equal when comparing the entire sample, the random selections from the sample show that the variability in number of patients who have a major bleed varies significantly, while the number of patients who are saved from having a stroke remains relatively constant. This indicates that there is a higher degree of imprecisions around the major bleed estimate, compared to the stroke risk estimate.

4.2.2 Violin Plot

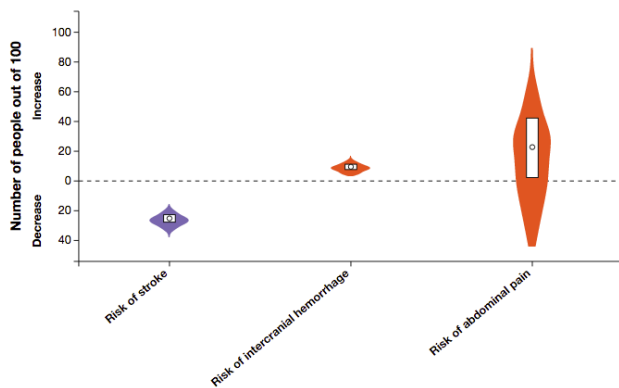


Fig 5. Violin plot: the risk of abdominal pain extends below the baseline zero-axis, indicating a potential benefit.

Violin plots are charts in which the underlying density-trace of the data is mirrored across the y-axis, with a box plot displayed inside the region, resulting in a shape similar to a violin. In terms of abstraction, the width of the violin at a specific y-coordinate corresponds to the likelihood that a given selection from the distribution will have that y-value. The minimum and maximum y-values of the violin encode the minimum and maximum probabilities of that attribute effect, according to the histogram results returned from D3. The y-positions of the maximum and minimum points in the box inside the violin respectively encode the 1st and 3rd quartile of the distribution. Finally, the circle mark y-position encodes the point estimate, which is either the mean or median depending on the toolbox setting.

Conceptually, violin plots are ideal as far as representing the degree of uncertainty in a risk estimate. By visualizing the entire

distribution, users can get an idea of likelihood across the entire range of possible values. Unfortunately, violin plots are not common, and many people unfamiliar with statistics may be confused as far as interpretation goes. We suggest that developers of DSTs think carefully as to whom their target population is, and determine whether violin plots might or might not be suitable as a visualization medium. In our case, when considering AF, we believe that the use of violins may be inappropriate, as the population is typically older than 60 years old and most likely unfamiliar with these types of charts.

Vital insight can be obtained from the analysis of violin plots, especially when the violin crosses the baseline axis. In Figure 5, the violin for the risk of abdominal pain has a significant portion of data below the baseline axis. This has serious implication as to the certainty of experiencing abdominal pain as a side effect. Theoretically, this indicates that the patient may see an improvement in abdominal pain as opposed to a worsening. Depending on how risk-seeking or risk-averse the patient is, this may significantly influence their decision making.

4.2.3 Gradient Plot

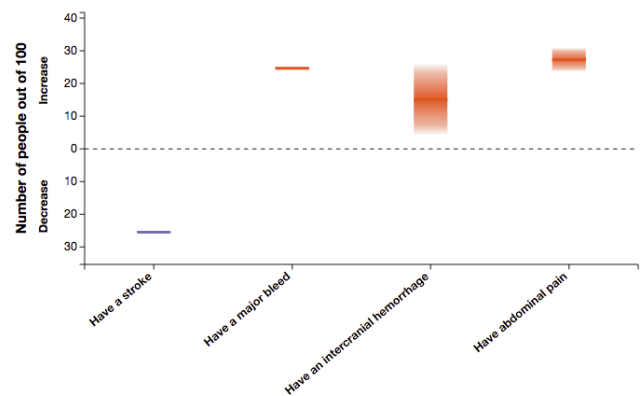


Fig 6. Gradient plot with large confidence interval around intracranial hemorrhage risk

The gradient plot is constructed as a basic scatter-plot, where a 3-pixel-high line represents each point on the scatter plot. In the case of extremely small confidence intervals, the actual point estimate can be occluded from the visualization when the point estimate isn't given a minimum height. The point estimate is therefore encoded by the midpoint y-position of the 3-pixel high line. A rectangle that is the height of one side of the confidence interval is then constructed, and placed on either side of the point estimate mark. The opacity at the y-position within the confidence interval encodes the probability of the true value being at that point, as per the underlying probability distribution.

While gradient charts better present the probability distribution in comparison to standard confidence intervals, they are not useful when the confidence intervals are extremely small. When that is the case, visualizations that actually present the underlying distribution may be more useful. However, in some cases, confidence intervals around point estimates may be very large, in which case they may provide useful insight to the patient. In Figure 6, the confidence interval around the risk of intracranial hemorrhage is large enough such that it overlaps that of having a major bleed, despite its point estimate being significantly lower. If a patient was highly anxious about having an intracranial hemorrhage, the large degree of uncertainty around that risk estimate may sway them away from warfarin treatment. This is an

example of a scenario where using gradient plots may be appropriate.

4.2.4 Interactivity

Through interaction, users can obtain valuable insight into some of the intricacies of the uncertainty surrounding their point estimates. While means and averages are useful as far making general statements goes, DST developers may wish to present patients with best-and-worst-case scenarios, as this a typical question asked in consult with physicians. In our tool, the best-and-worst-case buttons have a unique effect on the gradient bars in that the gradient bar disappears, and a standard bar chart appears showing the requested case. Due to the fact that our data is diverging and sitting on opposite sides of the baseline axis, the best-case represents the lower end of the confidence interval, where the risk of side-effects are minimized and chance of benefit is maximized. The reverse is true for the worst-case scenario. An example of what these might look like is shown in Figure 7.

Clicking on the best-case or worst-case button has an impact on the following visualizations:

- Gradient charts switch to bar charts displaying requested value at either upper or lower end of currently set confidence interval
- Isotype charts switch displays to values equal to gradient charts

Fig 7. Top left: Gradient chart transformed to bar chart in best case scenario. Bottom left: Gradient chart transformed to bar chart in worst case scenario. Right (top and bottom): Original gradient chart

Clicking on the mean or median radio button has an impact on the following visualizations:

- Gradient charts point estimates move along y axis to represent newly requested point estimate
- Isotype charts change value to requested point estimate
- Circle in box plot within violin moves to represent request point estimate

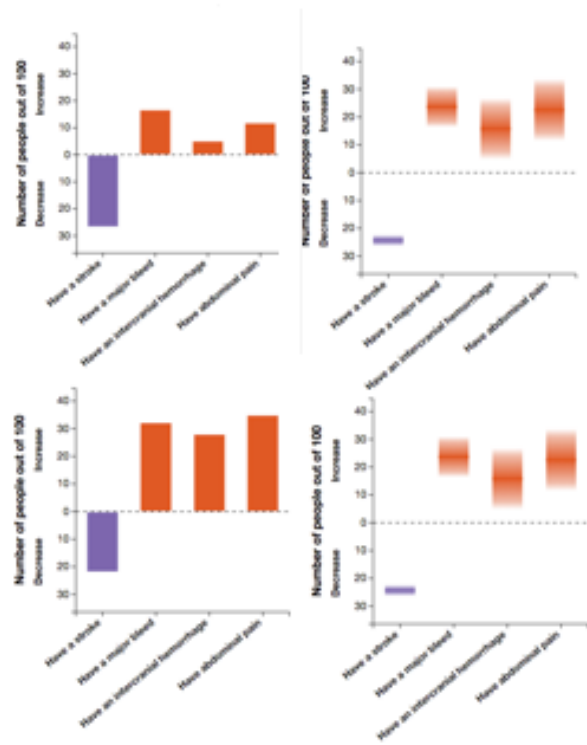
Clicking on the 95% or 99% degree of confidence radio button has an impact on the following visualizations:

- Height of confidence interval around point estimate on gradient chart adjusts
- If showing best-case or worst-case scenarios, gradient chart bars adjust height corresponding to requested degree of confidence
- If showing best-case or worst-case scenarios, isotype charts adjust number of people affected to requested degree of confidence

Clicking on the 100-person sample button has an impact on all of the visualizations, as all of the uncertainty parameters are recalculated based on the new sample, and charts are updated accordingly.

5 IMPLEMENTATION

To create our tool, we used the following libraries:



1. D3.js: Data-Driven Documents

This library was used to draw all charts in the visualization. Scaling and data-binding functions are used extensively.

2. AngularJS

This library was used to keep the structure of our project clean, along with simple two-way data-binding and reusable components through the use of directives.

3. Underscore.js

Underscore provided us with some useful functional paradigms, without needing to extend any built in objects

4. Bootstrap

This was used primarily for its grid-layout system, as it simplifies HTML page layouts. Some other features, such as wells and alerts, were also used.

5. NumPy

NumPy was used outside the visualization tool to generate our synthetic data, using the `numpy.random` distribution functions.

Of note is that all charts displayed in the visualization tool were developed from scratch as reusable components in AngularJS directives. No charting libraries were used to aid in implementation, apart from what was provided through D3. One of the primary challenges was drawing a proportioned outline of a person for use in the isotype charts, that could rescale as the window resized in order to keep the tool responsive. An additional relatively complex component was the violin plots requiring the use of D3s histogram and area functions, to actually interpolate between histogram scale positions and generate a violin.

Other tools used throughout the development of the tool were:

1. CoffeeScript

We chose to use CoffeeScript as an alternative to pure JavaScript, as it has a cleaner syntax and fewer curly-braces.

2. Sublime Text

We used Sublime Text as our text editor to manage the project.

3. Grunt

Used for its live-reload features in development, where files are watched for changes and the development server is updated to include those update files. Also used for building when it came time to deploy the software. This consists primarily of concatenating and minifying files.

4. Apache

An Ubuntu server running Apache was used to host the website, and a subdomain was created corresponding to the Apache configuration.

The tool was also designed with responsiveness in mind, and each individual visualization tapped into the window's resize event to scale itself to its container. When the window reached a width of less than 768 pixels, the elements stacked to keep the tool mobile friendly.

Additionally, the source code is available open-source and hosted on GitHub.

6 SCENARIO OF USE

In this scenario, Judy, a decision aid developer, has decided that she wants to create a decision aid for atrial fibrillation patients. She has conducted a large meta-analysis, and has a series of point estimates from hundreds of studies for a variety of attributes associated with atrial fibrillation treatment. This includes estimates for improvements in risk of stroke, alongside estimates for chances of side effects including major internal bleed, intercranial hemorrhage, and severe abdominal pain. She knows that some of these estimates have more uncertainty than others, but doesn't know the degree to which they differ, and isn't sure of the optimal way to present them.

She loads up our tool, and investigates the entire sample across all visualizations. She looks at the gradient chart, and sees that the confidence intervals around all 4 of the attributes are very small. She thinks that she must have been wrong, and that there is very little uncertainty at all in these estimates, but then notices the violin plots and realizes her mistake.

The violin for risk of major bleed extends from 5 out of 100 people to 50 out of 100 people, in a roughly Gaussian shape, while all other violins only cover a 5-person y-distance. To verify her concerns, she clicks the "Random 100 person sample" button a few times, and notices that in the major bleed attribute, the icon array changes drastically each time in the number of people who experience a major bleed, compared to the other attributes where the delta value remains relatively constant. She decides that, as this an older population, she should stick to using icon arrays as they have been shown to be effective as conveying risk in older populations. She decides to take screenshots of 3 different 100-person samples, along with a screenshot of the entire sample. She injects these images into her decision aid, with a textual explanation indicating that these are selections of 100 people from actual data, and this is what happened to them, so that patients get some understanding of how uncertain the risk estimates for risk of major bleed.

A patient, Joe, then goes through the tool, as he has just been diagnosed with atrial fibrillation and is deciding whether to take warfarin or not as treatment. Joe has been a big risk-taker for the majority of his life. He sees the risk estimates, and sees that in one 100-person sample, only 8 people had a major bleed. He thinks that maybe he'll get lucky, and be part of a sample like that. Based on his risk-seeking behavior, Joe decides that the pros of warfarin therapy outweigh the potential risks, and brings these thoughts up with his physician on their next consult.

7 DISCUSSION, CRITIQUE AND FUTURE WORK

By examining existing literature on how uncertainty can be encoded in risk visualizations, we developed a tool that can be used by DST developers to decide which visualizations are appropriate for their specific decision aid. The tool was interactive such that users could manipulate the uncertainty and identify how uncertainty can be encoded across various visualization techniques, including isotype arrays, gradient charts, and violin plots.

The primary issue that was encountered through the development of this project was with respect to acquiring the data. Our initial plan involved using clinical prediction models to compare real risk estimates. We invested a significant amount of time in researching these models, and settled on two commonly used ones for stroke and bleed risk, namely the CHA₂DS₂-VASC^[13] and HAS-BLED^[14] models. Unfortunately, while these models took as input some patient-related parameters and output specific risk estimates, there was no uncertainty reported with the estimate. Without any information regarding the uncertainty, we were unable to visualize anything apart from a point estimate. We debated making some assumptions as to what the underlying distribution looked like while using real model output values, but felt that this would be a misuse of the model and chose to fully synthesize our data instead. This has some implications as to the generalizability of our tool, as we are unsure how it might perform with real risk estimates. Our synthetic data artificially adds high levels of uncertainty to some attributes, and while we expect that this is something that exists in some disease contexts, including atrial fibrillation, we cannot be sure without future research.

Additionally, our model only presents second-order uncertainty, and opts to ignore the first order component. In the case of icon arrays, it would be relatively easy to add a first-order component. If 20 out of 100 patients experience some side-effect, we could randomly select 20 people on the isotype chart, as opposed to selecting the first 20 people. The issue here is that interpretation may be negatively influenced. If some interaction changes that point estimate from 20 to 30, and 30 new random people are shaded, the difference between 20 and 30 is less obvious when randomly distributed throughout the 100 people. Future work could add this level of uncertainty into the visualization, and see how users responded to its inclusion.

Another point of critique is how the visualizations actually move from the tool into DSTs. In the case of static charts, a simple screenshot or export of the underlying SVG element may be sufficient, but in the case where interaction is desired within the DST, some other means of export might be necessary. We can envision a system in which portable iframes or HTML packages can be exported, and then easily included within web-based DSTs.

Additional future work would involve improving the tool to enhance its generalizability. We can imagine a future iteration in which a backend is included, allowing users to add their own disease attributes and uploaded a dataset associated with that attribute. In such an implementation, the scope of the tool would not be limited to a single disease process.

Despite these limitations, we feel that our solution is a good first step at visualizing the various attributes associated with uncertainty. Due to the fact that uncertainty can't be represented by one single attribute, it is necessary to explore the various facets of uncertainty, including confidence intervals and the underlying probability distribution of the data. In doing so, users can get valuable insight into where uncertainty might exist, and therefore pass that on to patients in easy-to-understand ways.

ACKNOWLEDGMENTS

The author wishes to thank Tamara Munzner for her guidance in this project.

REFERENCES

- [1] O'Connor, A. M., et al. "Toward the 'tipping point': decision aids and informed patient choice." *Health Affairs* 26.3. 2007.
- [2] Elwyn, G., et al. "Developing a quality criteria framework for patient decision aids: online international Delphi consensus process." *Bmj* 333.7565. 2006.
- [3] Bansback, N., et al. "Communicating Uncertainty in Benefits and Harms: A Review of Patient Decision Support Interventions." *The Patient-Patient-Centered Outcomes Research*. 2016.
- [4] Brodlie, K., et al. "A Review of Uncertainty in Data Visualization." *Expanding the Frontiers of Visual Analytics and Visualization*, p. 81-109. 2012.
- [5] Johnson, B. B. and Slovic, P. "Lay views on uncertainty in environmental health risk assessment." *Journal of Risk Research*, p. 261-79. 1998.
- [6] Han, P. K. J., et al. "Communication of Uncertainty Regarding Individualized Cancer Risk Estimates: Effects and Influential Factors." *Medical Decision Making*, p. 354-66. 2010.
- [7] Correll, M. and Michael, G. "Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error." *IEEE Transactions on Visualization and Computer Graphics*, p. 2142-151, 2014.
- [8] Kao, D., et al. "Visualizing 2D probability distributions from EOS satellite image-derived data sets: a case study." *Proceedings Visualization. 2001*.
- [9] Hakone, A., et al. "PROACT: Iterative Design of a Patient-Centered Visualization for Effective Prostate Cancer Health Risk Communication." *IEEE Transactions on Visualization and Computer Graphics* 23.1. 2017.
- [10] Micallef, L., et al. "Assessing the effect of visualizations on bayesian reasoning through crowdsourcing." *IEEE Transactions on Visualization and Computer Graphics* 18.12. 2012.
- [11] Munzner, Tamara. *Visualization analysis and design*. CRC Press, 2014.
- [12] Galesic, M, et al. "Using icon arrays to communicate medical risks: overcoming low numeracy." *Health Psychology* 28.2. 2009.
- [13] Lip, G. Y. H., et al. "Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation." *Chest Journal* 137.2. 2010.
- [14] Pisters, R, et al. "A novel user-friendly score (HAS-BLED) to assess 1-year risk of major bleeding in patients with atrial fibrillation: the Euro Heart Survey." *Chest Journal* 138.5. 2010.