TITLE: Understanding children’s heart surgery data - a cross-disciplinary approach to co-develop a website

RUNNING HEAD: Developing a website on surgery outcomes

AUTHORS: Christina Pagel¹ PhD, Emily Jesper² BSc, Joanne Thomas² BSc, Emily Blackshaw³ BSc, Tim Rakow³ PhD, Mike Pearson⁴ MA, David Spiegelhalter⁴ PhD

¹. Clinical Operational Research Unit, University College London, 4 Taviton Street, London, WC1H 0BT, UK

². Sense about Science, 14a Clerkenwell Green, London EC1R 0DP, UK

³. Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King’s College London, Strand, London SE1 1UL, UK

⁴. Statistical Laboratory, Center for Mathematical Sciences, University of Cambridge, Wilberforce Rd, Cambridge CB3 0WB, UK

Corresponding author: Christina Pagel, Clinical Operational Research Unit, University College London, 4 Taviton St, London, WC1H 0BT. c.pagel@ucl.ac.uk Phone number: +1 781 428 0982

Keywords: Risk adjustment, Congenital heart disease, Public reporting of outcomes, co-production, public communication
ABSTRACT

Risk-adjusted survival statistics after children’s heart surgery are published annually in the UK. Interpreting these statistics is difficult and better resources about how to interpret survival data are needed.

Here we describe how a multi-disciplinary team of mathematicians, psychologists and a charity worked with parents of heart surgery children and other users to co-develop online resources to present survival outcomes.

Early and ongoing involvement of users was crucial and considerably changed the content, scope and look of the website, while the formal psychology experiments provided deeper insight.

The website http://childrensheartsurgery.info/ was launched in June 2016 to very positive reviews.
**Introduction**

Since 2000, all UK pediatric heart surgery centers have contributed data on every cardiac procedure to the National Congenital Heart Disease Audit (NCHDA) [1]. In 2013, the NCHDA began using the Partial Risk Adjustment in Surgery risk model (PRAiS) [2–4] to report risk-adjusted 30-day survival outcomes for participating hospitals. This triggered events resulting in the temporary suspension of surgery at one unit in 2013. The attending media scrutiny, impact on families and public anger [5–9], alongside previous and more recent media coverage on children’s heart surgery services [10–15] illustrate the need for resources to support appropriate interpretation of outcome data for families, journalists, clinicians and decision makers. As part of a project to update the PRAiS risk model (see accompanying papers [16–17]), we co-developed an explanatory website aimed at these diverse public and professional audiences (http://childrensheartsurgery.info/). It was a multi-disciplinary effort and was a transformative and valuable experience for the team (Table 1).

Following its launch in June 2016, a Lancet editorial [18] commended the website: “Many more areas of medicine requiring risk communication should take this initiative as a long overdue and most welcome example”. In this journal, Jacobs et al. have discussed the drive for public reporting in the US and highlighted the need for resources “that assist patients and families in correctly interpreting complex data” [19]. To encourage and help others to undertake similar ventures, here we describe our approach and the lessons we learned.

**Overall strategy**

The final output was aimed at two audiences:

1. Older patients and families of children who have had/will have heart surgery.
2. Other interested users including press officers and policy advisors for medical charities or professional bodies, the media, medical communicators, hospital family liaison services, and patient advocates.

An initial draft of content was followed by a process of iterative web development in response to user feedback. We involved both groups of target users from the outset, convening four 1.5-hour-long workshops for each group over a year. We aimed for four to six participants per workshop, with no-one attending more than once to ensure a fresh perspective. Participants received only minimal details about the project and were not required to read anything in advance. Workshops began with a brief background presentation and by establishing appropriate consent and permissions (e.g. recording). Next, participants were each given about 15 minutes to explore website material on a laptop (workshop 2 onwards). It was made clear to parents that they were free to leave if they did not feel comfortable to continue. There then followed a facilitated discussion, the focus of which varied between workshops as described below. We tested understanding of concepts and plain language explanations in all workshops. We identified where a concept had not been explained clearly and concepts which were especially difficult to understand. Participants were invited to provide feedback remotely on future iterations of the site (all accepted) and we incorporated their feedback on the near final web material.

At months 10 and months 12-15, we also shared web-content with the UK Children’s Heart Federation (CHF), specialists (including three pediatric cardiac surgeons, two cardiologists, two intensivists and two data experts) and representatives of the NCHDA (including senior clinicians) and incorporated their feedback. Starting 7 months into the project, the psychology team (TR, EB) used a mixed-methods approach to evaluate candidate components for the website. A summary of the strategy is shown in Figure 1.
Evolution of the site

Workshop participants

We held four rounds of two workshops, involving 15 participants in the parent workshops and 22 participants in the other workshops. Two team members facilitated each workshop, and two to three others attended as observers.

The feedback from the workshops and psychology experiments fundamentally influenced and changed almost every aspect of the website. Here we concentrate on two major themes: the development of the key messages and the data display.

Starting the web material development

The initial focus was on explaining the key table and graphic in the public NCHDA annual report [18] (Figures 2 and 3). The vertical axis in Figure 3 is the ratio of actual survival to predicted survival from the PRAiS risk model. If this value is greater (less) than one then survival at that hospital was higher (lower) than predicted. The white area gives the 95% prediction interval for this ratio – essentially the range within which we expect hospital outcomes to lie.

Members of the project team (CP, DS, MP, EJ) discussed potential content of the web material and plans for the first workshops. Fresh to the material, our web-programmer (MP) suggested that by rotating the chart (Figure 3) through 90°, it could become another column on the corresponding table (Figure 2), which might make the relationship between the table
and the graphic more transparent (Figure 4). We therefore decided to present participants in the first round of workshops with:

- Introductory text
- NCHDA table of results (Figure 2)
- NCHDA results graphic (Figure 3)
- A combined version of the table & graphic (Figure 4)
- Some basic explanatory text for the graphic
- Draft answers to example “Frequently asked questions” (FAQs)

Workshop round 1 (month 3)

Data display
Mathematicians DS and CP were both familiar with the way NCHDA presented the hospital outcomes (Figure 3). However, the value of the workshops became immediately apparent. Some workshop participants were drawn to the color regions in Figure 3 and interpreted it as a bar chart, which gave the (incorrect) impression that the hospitals on the horizontal axis are ranked by increasing survival. After discussion with participants, we decided to make the prediction interval colored and the outermost areas white, to focus attention on the dot and its predicted range. In both workshops, participants unanimously preferred the combined display (Figure 4), so we adopted it.

The workshops demonstrated that ratios are hard to understand and easy to misinterpret. DS and CP spent considerable time in both workshops explaining what the ratio of actual:predicted survival represented and why it was used instead of raw survival (we note this ratio is also used in the US [19]). The ratio is preferred by the NCHDA because every hospital has the same expected value of 1, which gives the graph in Figure 3 a common center line. As discussion progressed, CP and DS emphasised that hospitals should not be
compared directly to each other using their raw survival rates and that the key feature is whether the “dot” on the graph (representing the hospital) is within its predicted range. Discussing this aspect later, we realised that emphasising that hospitals should only be compared to their own predictions made it confusing to then transform hospital results to a ratio whose main benefit is to allow comparison between hospitals. We also realised that providing the exact predicted survival was inconsistent with emphasising that the predicted range was the important feature. We therefore decided to present actual survival for each hospital within its predicted range of survival, and not to provide the exact predicted survival rate.

Predicted survival
This created much discussion – participants asked when and how predictions were made. For instance, some people assumed that the hospital predicted its survival rates after knowing its actual survival, or that (analogous to predictions of poor performance in sports or education) hospitals with lower predictions were “worse”. We had tried to avoid using detail about the risk adjustment method in our explanations, instead using language like “predicted survival (%) is the percentage of operations where the child would survive at least 30 days after their operation estimated using previous national data about children similar to those cared for at that hospital.” This proved unhelpful, because different people interpreted this very differently. Parents in particular wanted to know more about how the survival was predicted. We realised that we would have to explain risk adjustment “up front” and learn how people interpreted the terms “predicted” or “expected”.

Workshop round 2 (months 8-10)

By the second “interested-user” workshop, we had developed three draft webpages: an introduction page (Figure 5), a data page (Figure 6) and an FAQ page. The introduction
page stated explicitly that the risk adjustment method was objective (based only on patient characteristics), and that raw survival rates should not be compared between hospitals.

For the data page, in addition to the changes identified from the first round of workshops, we included additional information available on clicking on a single hospital and by using “hover-overs”.

The second “interested-user” workshop (month 8)

Feedback was generally positive, with participants endorsing the site design, data display, and warnings against comparing hospitals by survival rates but they suggested that these warnings should be more prominent. Participants were reassured that all hospitals had high survival rates and suggested that this be emphasised. Most importantly, one participant commented that we emphasise that you should not compare hospitals and then present them in a table that invites comparison. The ensuing discussion prompted us to add a data view showing data individually by hospital.

The second parent workshop (month 10)

By this time, we had incorporated much of the feedback from the second interested-user workshop. Participants still wanted us to be clearer about what background information was absolutely essential and we discussed adding a “key points” section to the introduction.

We created a new “mapped data” page that allowed users to explore hospitals individually (Figure 7). We also provided links to the hospital website and any associated charity. There was still a separate page with the tabled data as in Figure 6.

The parents suggested that adding some explanation, plus links to relevant FAQs, alongside the individual hospital data displays would help.
First set of psychology experiments (months 10-11)

Three mixed-methods experiments explored how people understood and evaluated the prediction intervals and made subsequent judgements about hospitals, comparing the survival-ratio plot used in the NICOR Report (Figure 3) against the percentage-survival plot from the website (Figure 6). The findings confirmed our decision to use percentage-survival plots instead of ratio plots (e.g., accuracy of understanding for the outcome scale was 71% for the percentage plot compared to 41% for the ratio plot) but highlighted two key messages that we needed to emphasise even more:

- comparing hospitals’ survival rates to other hospitals’ survival rates or predicted ranges is inappropriate,
- the predicted range is based on an objective statistical formula and only reflects the hospital’s case mix.

The experiments and workshops also emphasised the importance of consistency in using (or implying) terms such as "luck", “chance”, “risk”, “uncertainty” or “probability”. We decided to always refer to predicted risk as "predicted chance of survival"; the placement of the survival dot in relation to its predicted range as “strong/some/no evidence that chances of survival at that hospital were different to that predicted using the formula" and inherent uncertainty in outcome as “unforeseeable factors”.

Workshop round 3 (month 11)

We focused on updating the site to be intuitive to navigate, to repeat key messages, to provide sufficient detail and to have a simpler home page (Figure 8). The introduction content was moved to a new page (“What, why, how?”) with a key points section at the top (Figure 9). Workshop feedback was now very positive but nonetheless participants suggested new content, wording revisions and layout changes.
Second set of psychology experiments (month 12)

An experiment found that if people viewed hospitals individually rather than alongside other hospitals, their interpretations of a hospital’s outcomes seemed more reliant on the prediction interval (and less reliant on other hospitals' outcomes). We therefore promoted the individual hospital view over the combined table/graph. We retained the combined view for transparency and its similarity to the NCHDA report presentation. However, we added text to emphasise that hospitals should not be compared to each other on survival rate. Further details of methods and data for the first two rounds of psychology experiments are available elsewhere [21].

Workshop round 4 (month 13)

With the web content and navigation nearly finalised, the project team concentrated on two explanatory animations. We worked with an external animation company but changed their usual development process by allowing time for user feedback at a relatively advanced stage of animation development. In detailed discussions, the parent participants suggested moving key sections from one draft animation (explaining the predicted range) to the other (explaining the data display) and then layering the placement of the animations within the site. This led us to substantially rework the storyboards and scripts.

Finalising the website (months 14-15) and evaluation

The workshops showed that the key points at the top of the “What/why/how” section were extremely valuable and should be prominent. We therefore illustrated the key points section of “What/why/how” (Figure 10) with images from the animations.
To determine whether the website provided “added value” over-and-above the NCHDA report [20], a small-scale randomised control trial compared people’s ability to answer questions about the audit data and its interpretation when they used either the NCHDA report (only), or the website (only), or both the report and website (together). Results showed that the website improves comprehension and understanding of the data, raising scores for these components by 0.75 and 1.25 standard deviations, respectively (see Appendix). The website offers three particular advantages over the report: it improves comprehension of the data plot, provides clearer information about the audit process triggered for outliers, and explains how often the survival rate will fall outside the predicted range by chance.

Lessons learned and summary

We set out to develop material to explain a single graph, but as we engaged with users, they helped us to develop messages about high survival rates, statistical methods, and avoiding inappropriate comparisons. By listening, we clearly communicated these messages and improved our presentation of the data. This early and continuous user engagement was vital, with each workshop improving and shaping the website. The multidisciplinary, multi-sector nature of our team was crucial to implementing this input (see also online articles from the perspective of a parent participant [22], Sense about Science [223,24] and the mathematicians [25,26]). The website, http://childrensheartsurgery.info/, was launched on 21st June 2016. It was very well received and quickly endorsed by major stakeholders (charities, clinical specialists, national audit, NHS Choices, Royal Colleges, leading medical journals) [18,27–31].

Building this website took considerable resources, including funding, people and time. It was extremely helpful to have an external partner guide the stakeholder involvement and act as neutral facilitators at the workshops. For the technical web development, responsiveness and speed in changing the web material was important, requiring a technical strategy for re-
working the website outside the norm for website development. Despite allowing ourselves 15 months, we were pressed for time – mainly because we underestimated the demands of implementing multiple iterations of feedback. For difficult topics such as survival outcomes, the parent workshops [22] in particular were humbling and vital for the team to understand the emotional aspect of the data for parents and inform development of sensitive communication. To summarise, there is no substitute for genuine co-production.

Acknowledgements

Funding
This project was funded by the UK National Institute for Health Research Health Services and Delivery Research Programme (project number 14/19/13).

The funder had no role in the writing of the manuscript or the decision to submit for publication.

Department of Health disclaimer
The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Health Services and Delivery Research Programme, NIHR, NHS or the UK Department of Health.

Author Contributions
Pagel (Operational Research) led the study and wrote the initial web content. All authors contributed to the final web content. Pagel wrote the first draft of the paper and all authors read and commented on the draft.

Jesper and Thomas (Sense about Science) organized and facilitated the workshops and planned and ran the website launch.
Rakow and Blackshaw (Psychology) ran the psychology experiments

Pearson (Web designer & producer of mathematical educational resources) built the website and data displays.

Spiegelhalter (Professor, Public Understanding of Risk) contributed to the website development, in particular the language around the communication of risk and worked closely with Pearson, Pagel and Sense about Science throughout.

**Conflict of Interest**

The time spent on this project for authors Pagel, Jesper, Thomas, Rakow, Blackshaw, Pearson was funded by the UK National Institute for Health Research Health Services and Delivery Research Programme (project number 14/19/13). Spiegelhalter provided his input at no cost. The authors declare no conflicts of interest.
References


[13] Perry K. Bristol Children’s Hospital “still in denial” about deaths, say parents. Telegraph.co.uk 20:03.


Table 1 – team reflections on co-developing the website:

Christina Pagel (mathematician): “People really valued having access to this information, presented in a clear and balanced way. I learned that accessibly presenting abstract and difficult concepts takes a huge amount of effort and listening. For anyone embarking on a similar project, do not underestimate the time needed and go for it - it was an eye-opening experience.”

David Spiegelhalter (risk communicator; statistician): “This has been a humbling and invaluable experience. I thought I knew something about communicating statistics, but sitting listening to enthusiastic users struggling to understand concepts made me realise my inadequacy. If we want to genuinely communicate statistical evidence, I am now utterly convinced that users have to be involved from the very start.”

Tim Rakow (experimental psychologist): “Collaboration has been key to the success of this project. I believe we have developed something that should allow people to engage with what would otherwise be fairly opaque information.”

Emily Jesper (workshop facilitator; science communicator): “It was fantastic to feel that researchers involved in the project were so responsive to the user feedback and careful to consider how to sensitively communicate and not afraid to ask participants about issues they were stuck on. It highlighted why involving the audience early cannot be underestimated, and we urge other researchers to adopt this approach.”
FIGURES

FIGURE CAPTIONS

Figure 1 - summary of the development plan for the website. CHF = Children’s Heart Federation. NCHDA = National Congenital Heart Disease Audit.

Figure 2 - Table of outcomes taken from the NCHDA 2010-2013 report which accompanies Figure 3

Figure 3 - example of the key NICOR output, taken from the 2010-2013 annual report.

Figure 4 – new horizontal display of the graphic alongside the table content.

Figure 5. Our first attempt at the introduction page (showing just the top and the bottom of the page).

Figure 6 - our first attempt at the data page

Figure 7 - the mapped data tab showing the display for a specific hospital

Figure 8 - the new, simpler home page developed for the third set of workshops. This is its final version from the live website.

Figure 9 - the new "what, why, how" page developed for the third set of workshops

Figure 10 - Final set of "key points" at the top of the "What/Why/How" webpage
A small randomised control trial presented 35 participants with: 11 questions testing comprehension of the prediction interval plot (e.g., Where would the dot be if there was some evidence that the hospital’s survival rate was lower than predicted?); and 15 questions testing understanding the monitoring process and the influences on the survival data (e.g., How often is survival rate outside predicted range by chance?). While answering these questions, participants could access either (1) the NCHDA Report for these data, or (2) the explanatory website, or (3) both the NCHDA Report and the website. Questions were scored for accuracy; these scores were combined to create percentage accuracy scores for comprehension, understanding and an overall score (combining both elements).

Mean comprehension accuracy was significantly greater for those with access to the web-tool (M = 90.5%, SD = 16.1%) compared to those given only the report (M = 78.8%, SD = 13.0%), t(33) = 2.17, p = .037, 95%CI M\text{difference} 0.8-22.7%, d = 0.77 (large effect). Mean understanding score was significantly higher among those with access to the web-tool (M = 53.8%, SD = 17.7%) than among those only given access to the report (M = 31.8%, SD = 17.2%), t(33) = 3.52, p = .001, 95%CI M\text{difference} 9.3-34.7%, d = 1.25 (large effect). Mean overall survey score was significantly higher among those with access to the web-tool (M = 69.3%, SD = 13.7%) than among those only given access to the report (M = 51.7%, SD = 11.3%), t(33) = 3.83, p = .001, 95%CI M\text{difference} 8.3-27.0%, d = 1.36 (large effect).

For more details on the experiments please see Rakow et al. [21].