JODIE MCVERNON
NICHOLAS GEARD
JAMES MCCAW
MATT DUCKHAM
SUSANNE BLEISCH
FERNANDO J MARTIN-SANCHEZ
GUILLERMO LOPEZ-CAMPOS

IMPROVING THE COMMUNICATION OF SCIENTIFIC RESULTS TO POLICY MAKERS

28 JULY 2014

MELBOURNE SCHOOL OF GOVERNMENT
INCUBATOR RESEARCH PROJECT GRANT SCHEME FINAL REPORT

DOI:10.5281/ZENODO.15331321

1 Introduction

Mathematical models are increasingly used to help policy makers decide among alternative strategies. For example, in the domain of population health, mathematical models are used to understand the potential impacts of emerging infectious diseases, and the potential responses of those diseases to a range of policy interventions.

However, the mathematical models are complex. Uncertainty is endemic in any such predictive model. Further, even sophisticated models necessarily incorporate a number of simplifying assumptions and include input parameters which have implications for the correct interpretation of model outputs. The outputs from these models are also frequently multidimensional, with complex data covering a range of aspects of the phenomenon.

Thus, accurate and comprehensible communication of model outputs is challenging, but essential for enabling policy makers to make informed and effective decisions.

This incubator project sought to understand better the role of scientific model outputs in the policy process, in particular in the context of the health domain. By looking in parallel at the process of policy formulation and the form of the scientific reports used in that process, the project was successful in identifying important recommendations that can improve the communication of scientific results to policy makers. The project also highlights the potential for substantial further gains to be made from future projects in this area.

1.1 Research projects and reports used

The methods and guidelines developed in this project applied to two different scientific projects, which aim to advise the Australian Government Department of Health, Office of Health Protection, on two different issues. The projects/reports are:

Report 1: "Development of options on how to define the concept of pandemic impact for Australian purposes" which reported on simulation modeling and was written by J. McVernon and J. McCaw, 26th June 2012 [1]


DEVELOPMENT OF OPTIONS ON HOW TO DEFINE THE CONCEPT OF PANDEMIC IMPACT FOR AUSTRALIAN PURPOSES

REPORT ON SIMULATION MODELLING

Reporting Deliverable 3 to the Office of Health Protection

Australian Government Department of Health and Ageing

A/Prof Jodie McVernon, Dr James McCaw

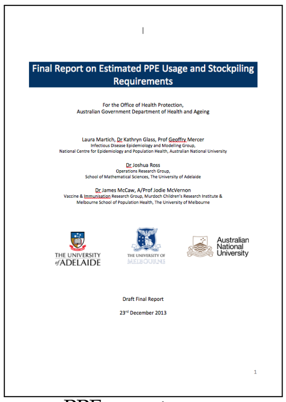
Vaccine & Immunisation Research Group
Murdoch Childrens Research Institute &
Melbourne School of Population Health, The University of Melbourne

26th June 2012

Figure 1.1: Pandemic impact report

This report was already finalized and submitted by the time the project started. It was used as an example report and analyzed for the type of communication used, especially in regard to data graphics.

Report 2: "Final Report on Estimated PPE Usage and Stockpiling Requirements" which reported outcomes of mathematical models and was written by L. Martich, K. Glass, G. Mercer, J. Ross, J. McCaw and J. McVernon, 23rd December 2013 [2]
This report was finalized and submitted during the project. The authors made use of the knowledge from the consultation workshop in Canberra to improve the structure of the report, especially the writing. The finalized report was then used to apply the recommendations in this report and to create improved data graphics able to illustrate the management summary of the report.



Final Report on Estimated PPE Usage and Stockpiling Requirements

For the Office of Health Protection,
Australian Government Department of Health and Ageing

Laura Martich, Dr Kathryn Glass, Prof Geoffry Mercer
Infectious Disease Epidemiology and Modelling Group,
National Centre for Epidemiology and Population Health, Australian National University

Dr Joshua Ross
Operations Research Group,
School of Mathematical Sciences, The University of Adelaide

Dr James McCaw, A/Prof Jodie McVernon
Vaccine & Immunisation Research Group, Murdoch Children's Research Institute &
Melbourne School of Population Health, The University of Melbourne

THE UNIVERSITY of ADELAIDE
THE UNIVERSITY OF MELBOURNE
Australian National University

Draft Final Report

23rd December 2013

1

Figure 1.2: PPE report

10.5281/zenodo.15331321

2 Methods

The study followed a process that is outlined through following 7 steps:

1. Evaluation of existing communication through analysis of submitted (pandemic impact) report.
2. Requirements analysis to analyze the commonalities and differences in the communication requirements of the two parties: a) the scientists concerning what they need to communicate; and b) in consultation with the policy advisers concerning what their information needs are.
3. Development of recommendations for improved data graphics, plus advice on how to apply the recommendations in the data graphic design process.
4. Redevelopment of visual elements in the PPE report based on recommendations for the pandemic impact report.
5. Consultation with policy advisers regarding improved (visual) communication of scientific results.
6. Application of the recommendations and the data graphic design process to create scientific data graphics for a new report.
7. Summary of all feedback and improvement of the recommendations and the advice on the data graphic design process.

2.1 Report analysis (pandemic impact)

An existing report [1], which was used to communicate scientific results of pandemic impact simulations to the Office of Health Protection (Australian Government Department of Health and Ageing), was analyzed. The main focus of this analysis was on the data graphics used, and their appropriateness given the information communicated and the audience. However, the report was also used to learn generally about the style used to communicate research results in this specific area.

The graphical and non-graphical representations of data used to communicate scientific findings in the report included tables, boxplots, tornado plots, and heatmaps. Examples of these include:

Boxplots

Boxplots are a common graphical representation of results originating from repeated measurements or in this case from a series of simulations. Boxplots show the full range of the data and indicate key values such as the median. In the report, a trend line was added to the boxplots (Figure 2.1). Problems encountered include the use of hard-to-interpret abbreviations as axes labels. Additionally, the amount of information contained in the graphic may be overwhelming for audiences not already familiar with the topic and this specialized representation type.

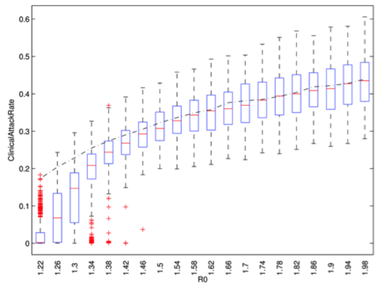


Figure 2.1: Boxplot example

Tornado plots

Tornado plots are a variation of bar charts that show positive and negative influences of parameters in the simulations in this example (Figure 2.2). Bar charts are common, easy to read, and allow easy comparison between values. The problems of tornado plots and bar charts include the use of abbreviations and symbols for the axes labels.

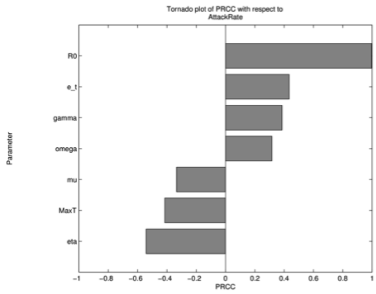


Figure 2.2: Tornado plot

Heatmaps

The heatmap shows pandemic attack rates dependent on two parameters (the two axes, Figure 2.3). Additionally, these representations are used to show percentage of change. The health advisers reported that they found these representations difficult to read as it was unfamiliar to them. Generally, the use of colors to show magnitude is problematic as we do not perceive distances between colors constantly, and the colors are not perceptually ordered to give a clear value message ("more" of something) [3].

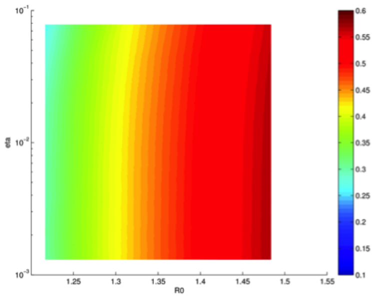


Figure 2.3: Heatmap

Overview representations

In the executive summary, a schematic chart was used where the graphic positions of different pandemic scenarios were plotted on two axes: severity and transmissibility (Figure 2.4). Severity and transmissibility are two key influencing factors for pandemic impact. This graphic was much liked by the policy advisers and made it through to further summaries of the work. The graphic is simple and clear and importantly allows the scenarios to be related to experience, through the blue bubbles which indicated the approximate position of previous pandemics. The use of colors and shading could be improved to make the representation even more readable.

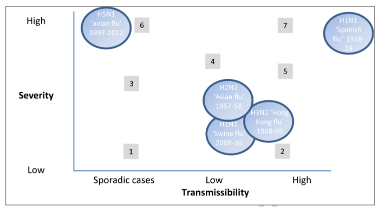


Figure 2.4: High-level graphic

Tabular representations

Positive comments made about the use of tables (Figure 2.5) to represent results coincide with the expressed desire of health professionals for communication in exact numbers. Uncertainty is

Scenario (see figure 1)	Severity	Transmissibility	Average Attack rate	Average Clinical attack rate
1	Low	Low	25-40%	0-5%
2	Low	High	>50%	0-5%
3	Moderate	Low	25-40%	0-10%
4	Moderate	Moderate	40-50%	5-12%
5	Moderate	High	>50%	8-12%
6	High	Low	25-40%	10-17%
7	High	High	>50%	15-25%

Figure 2.5: Tables example

included through the use of ranges. The table presents the values in a structured way in relation to the seven scenarios defined through severity and transmissibility as detailed in the representation above.

2.2 *Requirement analysis*

A requirements analysis meeting was undertaken with the scientists working on the existing report [1], but also the new reports [2]. A second requirements analysis meeting with the health advisers of the Office of Health Protection, Australian Government Department of Health, was used to learn about their needs in receiving communication of research results from scientists.

Each meeting lasted for about 60 minutes, with detailed notes made during the meeting. Participants were encouraged to talk freely about their information needs, the current processes of policy making, and the communication of scientific results in this topic area. The notes were subsequently analyzed to find key areas of overlap in goals, or important mismatches in communication, with a specific focus on the creation and/or improvement of data graphics.

10.5281/zenodo.15331321

3 Results

The issues identified through the requirements analysis showed a large overlap between the information the scientists want to communicate and the information the policy advisers want to receive. However, the discussions also showed that there is still room for improvement in the communication of scientific results to policy makers.

Additionally, some issues identified by both sides do not have similar levels of importance, or reasons for importance, when viewed from the scientist's or policy-maker's perspectives. From the discussions, six key requirements and constraints for visual communication were identified, described in sections 3.1–3.6 below. The first four issues are relatively general communication issues with some relevance for the development of visual communication. The additional two issues are more visualization-specific. The numbers (①, ②, ③, ④, and ⑤) used in the "recommendations" section of each description link to respective parts of section 4, where the different levels of recommendations for improved visual communication are laid out in more detail.

3.1 Layering of information

Origin: The original reports contain some layering of information, primarily in the summary, the main body of the report, and potentially any appendices. However, the importance of layering of information was emphasized in particular from the policy adviser side, where this aspect is crucial for further communication of the scientific results to decision makers.

Details: Information from the models needs to be represented and summarized at different levels of detail, including: key messages, executive summary, scientific report, and technical details. Layering is important as the information is used in different ways by different audiences. The audiences have different amounts of time available to assimilate the information.

Recommendations:

- Follow the general advice for representations as defined in ① throughout all layers.

- Use the same visual elements ② in each layer of information, but vary the complexity and/or granularity of the information content.
- Additionally, the different layers could be linked to allow following up on details when viewing more summarized levels. ⑤

3.2 *Providing decision support*

Origin: The scientific reports are generated to provide decision support. As the scientists themselves are aware "This work can't be done without an opinion." However, decision support must be based on scientific results and the layered structure of the report (described above) should help the user to access the technical details. The need for decision support was mentioned by both the scientists and the policy advisers.

Details: One aim of communicating the scientific results is to provide decision support. Communication needs to be based on evidence; potentially show critical decision points; and detail outcomes and likelihoods of success including consequences. Additionally, reports must include influencing contextual factors, the implications of uncertainty of the information, as well as different views or even opinions.

Recommendations:

- Decision support can be achieved through "telling the story," potentially through adding narratives or text to data graphics ③.
- Decisions are also supported by better linking between the different levels of layered information, to allow policy makers to follow up on details ⑤.
- Critical decision points or outcomes are more difficult to integrate in (simple) representations ⑤.
- Influencing factors could be added through story-telling ③ or different visual elements ②.

3.3 *Hierarchy of importance*

Origin: The hierarchy of importance was not an explicitly stated requirement. However, through discussion in both groups it became clear how important it is to frame and structure the information and to provide some reference and options to relate the presented information to experience and context.

Details: Not all aspects of the scientific results may be similarly important in communicating the results in a specific scenario.

The most important aspects could be used to frame or structure the presentation of the results or information. In some cases some variables explain much more variation than other variables. For example, in the pandemic impact study the aspects severity and transmissibility are more important than others.

Recommendations:

- Use the most important aspects to structure the information and potentially to provide a reference frame or context for the information ④. For example, use scenarios or levels of information (such as transmissibility and severity) as structuring axes.
- The information may then be related to experience (e.g., previous pandemic influenza scenarios), could define the granularity of information (e.g., through levels or extremes, such as best/worst), and could show different input options and/or different outcomes.

3.4 *Influencing factors*

Origin: The importance of influencing factors was not explicitly stated in a compact way, but aspects of this issue came up several times through both sides' discussions.

Details: Factors like familiarity, simplicity, memorability, trust, preferences, and culture were frequently mentioned as being important. Using simple and familiar ways to represent information reduces the time required to assimilate it. Additionally, using familiar representations can add trust and thus better acceptance to the communication process. Previous studies have shown that preference (or dislike) of a representation type can influence performance.

Recommendations:

- Use a selection of visual elements from a defined range ② to communicate the information. This allows familiarity but also the accommodation of different preferences.
- The visual elements should be simple, crisp, understandable, memorable, and not obscure the message through decoration. ①
- Trust can potentially be enhanced through adding more details to the data graphic—telling a story ③—as well as linking to further details ⑤. Additionally, providing a structure that allows users to relate to experience could be helpful ④.

3.5 *Information types*

Origin: The different information and data types mentioned in the two discussion groups did mostly overlap.

Details: The policy advisers stated a preference for quantitative information, especially exact numbers. However, they also understand how important it is to know how certain or uncertain those exact numbers are. Specifically, they are interested in the influence the uncertainty can have on a decision, i.e., what is the risk associated with a decision? For the scientists the communication of uncertainty associated with exact numbers is crucial as the simulations and modeling process do not produce absolute values as output. The types of information to be communicated include: quantitative data, qualitative data, uncertainty information, time and duration, as well as change.

Recommendations:

- Combine selected visual elements ② that are able to communicate ranges and uncertainty with numbers ③.
- Levels of information ④ could help in communicating numbers while still conveying associated uncertainty.
- Specifically designed *infographics* ③ may be able to communicate decision points in time as well as risks associated with certain decisions based on the uncertainty associated with the data. ⑤

3.6 Technical requirements

Origin: The technical requirements were mainly elicited from the discussion with the policy advisers. The preferences for printable files may be slowly changing as more and more decision makers use tables or smartphones to access information and reports.

Details: The resulting data graphics should be printable in some way (e.g., screenshots of interactions) as many decision makers still prefer hard copy print-outs. There is potential to include data graphics in the documents (e.g., PDF) or have them as separate files. However, separate files may not be looked at. Graphics should be small in size as they are generally distributed via email. Data graphics should be useful for presentations and so projections. Interactive interfaces remain an option. However, there may not be enough time to play with interactions. Further, too many options may become confusing.

Recommendations:

- Follow the general advice for representations which includes a section on technical requirements such as small file sizes or suitable colors for different media or audiences ①.

3.7 Further discussed aspects

A further aspect that was discussed with some intensity, especially in the consultation with the policy advisers, was the issue of timing.

It is often crucial to provide the right information at the right time and according to the given structure, such as duration or format. This is not specifically a visualization problem, but structuring information at different levels and providing data graphics associated with these different levels may be helpful.

Finally, it was mentioned that providing decision makers with scientific results requires a close interaction between the scientists and policy advisers. It is crucial, for example, to define the exact questions to be answered through research, but also to understand the general context for the request. Close collaboration and feedback seems essential.

10.5281/zenodo.15331321

4 Recommendations

This section introduces three, increasingly complex levels of advice (items ①, ② and ③ below) to improve visual communication of information to decision makers. The first level ① comprises general advice to improve any representation, but with a specific focus on the requirements as defined earlier (see section 3). The second level ② defines a number of potentially useful visual elements to be used to show proportions or percentage values as well as discussing the layout or alignment of information along sensible axes. The third level ③ is concerned with the creation of infographics and is likely to be the most complex to implement.

Bringing together these three levels of advice, the fourth item ④ tries to outline a process of how our three levels (items ①, ② and ③) could be used to create improved graphics. This process is similar to the process of report writing in general. Thus, report writing and creating representations should be parallel processes.

Finally, the fifth item ⑤ takes up the issue of layering information in representations, which was not a detailed focus of this project. There is need for further research on this issue.

4.1 General advice for representations ①

The analysis of the graphics of the existing report on Pandemic Impact (section 1.1, Report 1), together with results of the requirements analysis (most especially results on technical details) initiated the compilation of general advice for representations to be used in conjunction with reports communicating the outcomes of scientific results to policy advisers. The aim was to generate a checklist (see Appendix 5.1) that would ensure that representations in reports would adhere to some minimal requirements in three thematic areas.

Our first area of advice is concerned with producing comprehensible overview representations. The representations (specifically the ones that are used as summaries) should be self-contained and standalone. It is possible that the representations may be removed from the report, and be used in meetings without captions or accompanying text. This facility requires descriptive titles, clear and legible labeling, and legends without abbreviations and symbols, or where there are explained in the legend. Despite the need to be self-contained, the representations may best be included in the report and also explained further in accompanying text. However, these summary representations should not be dependent on additional text or explanation in the report to be understood.

Second, technical considerations need to be taken into account in generating data graphics for scientific reports. The data graphics are frequently used in presentations, printed out, or sent by email. This requires file sizes to be small, and requires selective and considered use of colors. Small files can be achieved either through vector formats or by using the Portable Vector Graphics (png) file format as well as compression techniques. It is important to ensure data graphics continue to convey the correct message in black and white or grey scale. Sparing use of color will help to ensure a data graphic reproduces well in grey scale reproduction. Aiming for a maximum resolution of 1024 × 768 pixels also helps to ensure the graphics are rendered correctly with a variety of office projectors and keeps file size small.

Third, data graphics can be improved by following [4] advice for graphical excellence, graphical integrity, data-ink maximization, and the reduction of chart junk. It is important to make sure that data graphics allow comparison of information, by aligning data to the same scale, especially when placed side-by-side. Generally, it improves graphics to avoid redundant encoding of information, for example, by using color and height to show the same information. However, some highlighting or annotations may be added to emphasize key messages or the narrative flow of a story being told with the representation.

4.2 Visual elements for the communication of key content ②

A study by Cleveland & McGill [5] has shown which graphical elements work best for the accurate communication of numerical values. For example, length is more accurately perceived than angles, which in turn outperform area perception. However, as indicated in the requirements analysis (section 3.5) it is important to communicate not only the values but also the associated uncertainty stemming from a range of factors.

The subsections that follow show and discuss examples of visual elements that specifically can be used to show parts of a whole or proportions, such as percentage values. Most of them also include some form of visual representation of associated uncertainty. Not all of them work similarly well and, as learned through consultation, also peoples' different preferences or the intended audiences should be taken into account when selecting visual elements.

Another important aspect is the layout of the chosen visual elements. The requirement analysis showed a strong preferences for structured layouts such as tables (see Figure 2.5) or alignment of information with levels of factors or previous experiences (see the "bubble graphic" in Figure 2.4). Thus, it is recommended to think carefully if one of these principles can be used to arrange the chosen visual elements. Doing so has the effect of "framing" the information with context (e.g., key influencing factors), facilitating comparison (e.g., comparing different values along the axes levels) and highlighting the relationship to (potentially simplified) concepts (e.g., levels of severity or weeks into the pandemic) and previous experience (e.g., previous pandemic scenarios).

4.2.1 *Randomized dots*

Randomized dots communicate ordinal information rather than precise numerical magnitudes. Further research is needed to define how precisely humans can perceive the levels and what the range of values is that is communicated by a specific randomized dot representations.

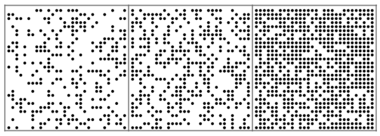


Figure 4.1: Randomized dots

4.2.2 *Randomized dots with uncertainty*

Adding uncertainty to the random dot plots does not seem to work well. The resulting impact is rather chaotic and difficult to interpret. Further research might be needed to improve this representation. However, a number of unfavorable comments in the discussions indicate that this type of representation should rather not be used.

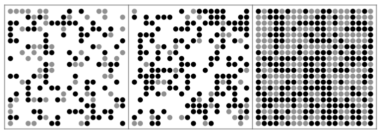


Figure 4.2: Randomized dots with uncertainty

4.2.3 *Regular dots with uncertainty*

An alternative option uses a regular array of the black dots, representing uncertainty using grey dots. This allows the perception of quite precise numerical values. Additionally, it indicates the level of uncertainty. For example, the middle presentation in Figure 4.3 has the lowest associated uncertainty when compared with the other two representations in this example.

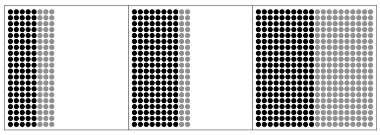


Figure 4.3: Regular dots with uncertainty

4.2.4 *Doughnut*

Doughnut graphs can communicate proportions. Uncertainty is indicated through the grey area. However, the grey tends to be perceived as a different value, instead of uncertainty associated to black area. Frames can help the interpretation of the values.

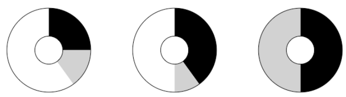


Figure 4.4: Doughnut graphs

4.2.5 *Doughnuts with blurring*

Adding uncertainty as through blurring can help address the issue above: incorrectly perceiving the "grey" area as a different, independent value.

Figure 4.5: Doughnut graphs with blurring

4.2.6 *Pie charts*

Figure 4.6: Pie charts

The circle uses the same concept as the doughnuts, but is visually heavier. Its effect may be improved without a reference frame. The grey area representing uncertainty is similarly difficult to relate to the black area as in the doughnut representation.

4.2.7 *Blurred bars*

Figure 4.7: Blurred bars

Bars are a familiar data graphic (an important factor as discussed in 3.4). Additionally, bar length is more easily compared than angles in doughnuts or circles [5]. However, a frame indicating the whole is needed to perceive it as a parts of a whole or proportion representation unlike to circles or donuts which naturally represent a whole and proportions. Uncertainty can be added as blurring or alternatively as a grey area.

4.3 *Illustrate the story and/or key messages: Infographics* ③

An "infographic" is "a visual representation of information or data, e.g., as a chart or diagram." Inforgraphics originated in the 1960s, a portmanteau of "information" and "graphic." In recent years, the term "infographic" has come to denote a type of visualization that uses a carefully designed combination of graphical elements, layout, and text to illustrate, highlight, or summarize a topic or to "tell a story." As an illustration, Figures 4.8–4.10 gives three examples of recent infographics on health topics. Many more examples may be found searching using the keyword "infographics." It is important to note that infographics often are usually tailor-made for a specific purpose and audience. That specification likely explains some of the choices of graphics, text, and colors. When designed carefully, infographics can be highly effective in the stand-alone communication of a summary or story.

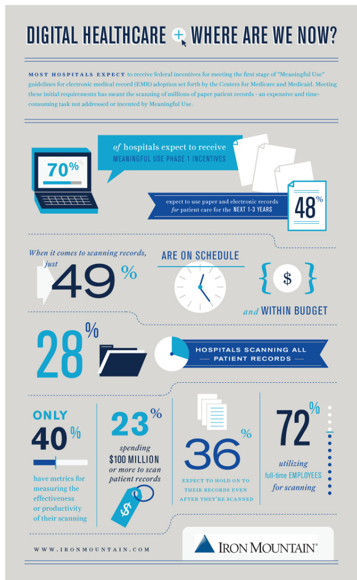


Figure 4.8: Example health infographic, after [6]

The process for creating infographics (and representations for communication in general) is sketched in item ④ below.

4.4 *The process of creating representations* ④

The process of creating representations for communication of information, including infographics, can benefit from close alignment with the report writing process. The following steps are aimed at guiding through the authors of scientific health reports through the process.

1. *Define the key message(s) to communicate*. This decision requires domain knowledge and may benefit from close collaboration between scientists and their audience. For all following steps, consider for each choice or decision whether the graphics communicate these key messages without distorting them. Be sure not to overload each data graphic. Potentially summarize or

aggregate the information to support the key messages and/or consider creating more than one representation for different aspects of each message or at different levels of details.

2. *Determine whether there is a "natural" structure to the information to be presented*. If there is no natural structure, it may be advantageous to use the most important influencing factors (such as the severity and transmissibility in the pandemic impact study [1]) as structure. Alternatively, one might decide to structure data graphics along one or more storylines or relevant example scenarios.

3. *Select suitable visual elements for encoding the information*. In choosing suitable visual elements it may be necessary to experiment with different options, or let the audience choose from elements to suit their preferences. Make sure the visual elements match the purpose of the data graphic and support the communication of the key messages.

4. *Follow a storyline* For infographics, chose simple graphical elements that illustrate a scenario. Carefully chose a suitable layout. Tell the story or highlight the key messages through suitable text and markings such as color, frames, or arrows. You may want to use familiar/memorable elements such as relevant pictograms instead of or in addition to labels.

5. *Refine the representations* Review, critique, and refine the representations through the general advice (item ① above) for improving representations.

6. *Layer information* For the layering of information (see also item ⑤ below), it may be useful to consider using the same structure or layout of the information but with a finer or coarser granularity respectively. Additionally, the same or similar visual elements could be used to communicate more or less details depending on the level the representation is intended to communicate at. A few lessons in this regard may be learned from cartographic generalization (e.g. [8]).

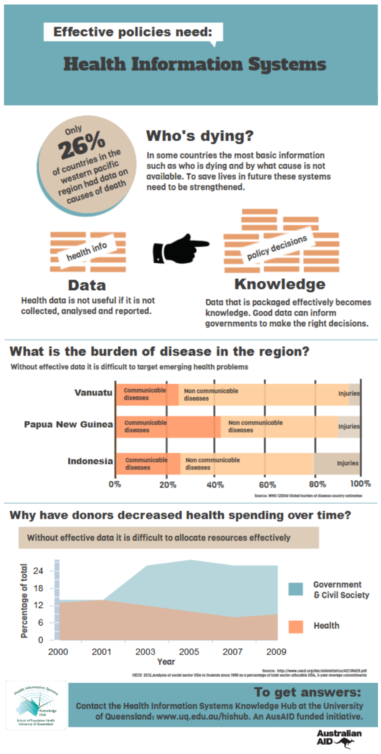


Figure 4.9: Example health infographic, after [7]

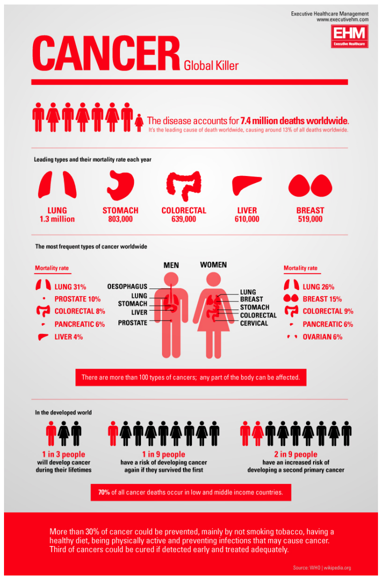


Figure 4.10: Example health infographic, after [9]

4.5 *Unsolved problems requiring further research* ⑤

One aspect of report and data graphic preparation that was revealed by the requirements analysis concerned the layering of information at different levels of detail in the reports. While this finding was focused primarily on the written report, a similar concept could be adopted for layering information in data graphics. Open questions include: how can visual representations at different levels be (interactively) linked? One option is the use of annotations and links that could be switched on or off based on user preferences. These links would then enable navigation between the different layers of information. A report might start with a summary data graphic of the content, in order to get an overview. Upon request, users would then be able dig deeper into the data and evidence by following links that reveal further data graphics with more details and data. As discussed above, it may be useful to use the same or similar visual elements but at different levels of aggregation.

Such possibilities lead to the consideration of interactive or animated report elements. Interaction demands additional cognitive load of users. Consequently, interaction would need to be carefully designed to support specific information and activities without overloading already time-poor policy advisers. The options that interaction could provide and their effective design is not discussed in this project, but is an important avenue for future research in this area.

10.5281/zenodo.15331321

5 Applications

The recommendations as detailed in section 4 were applied to the design of data graphics for the personal protective equipment stockpiling report [2]. Two example original images in the summary are shown in Figures 5.1 and 5.2. The goal of redesigning these data graphics was to create a standalone representation that summarizes the key messages of the report. Although a few infographic elements for highlighting were used, the aim was not specifically to create an infographic. The resulting data graphic is shown in Figure 5.3.

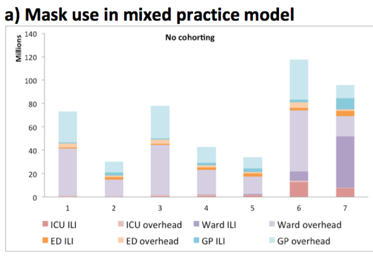


Figure 5.1: Example original PPE data graphic [2].

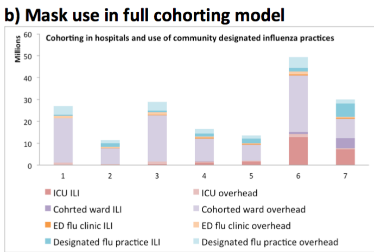


Figure 5.2: Example original PPE data graphic [2].

However, even without creating completely new data graphics the original graphics (as shown in Figures 5.1 and 5.2) could still be improved by working through the checklist in Appendix 5.1. The following paragraphs show how the advice in section 4 part ④ was applied to create the summary data graphic of the key messages from the report [2], shown in Figure 5.3.

1. *Defining the key messages and identifying supporting data*: The key messages for the data graphic were taken from the section "1. Key Findings" of the PPE Stockpiling report [2]. From the summary of the report it became clear that the two existing data graphics (see Figures 5.1, 5.2) were intended to support the communication of the key messages. Thus, those two data graphics were used as a starting point. By focusing on simple data graphics, and taking care not overload the reader with information, we checked how much of the information in the existing data graphics related directly to the key findings. Critical analysis of these data graphics revealed that the key findings are related a. to medical staff use of PPE dynamically associated with case presentation (ILI); and b. overhead use of PPE for general protection of hospital and clinical staff regardless of direct patient contact (Overhead). The more detailed subcategorisations (see Figure 5.1, 5.2) are arguably not as important as these key findings. It was thus decided to aggregate the data contained in the original figures to only show the two categories, ILI and Overhead. Using these two categories directly supports the communication of the key findings, which are also formulated at this level of granularity.

2. *Defining a structure/layout*: Figures 5.1 and 5.2 contain information related to seven pandemic scenarios. Those seven scenarios were laid out in an earlier report [1] along the two axes of severity and transmissibility (see the "bubble graphic" in Figure 2.4). That graphic received positive feedback from the policy makers. We thus decided to reuse the layout of that graphic (the axes of transmissibility and severity subdivided in three categories low, medium, high) as the overall layout to show the key findings of the stockpiling report. The selected visual elements (see following step) were laid out along those axes.

3. *Selection of visual elements*: The key messages talk about comparing the two categories, ILI and Overhead, as well as comparing between two models (standard and cohorting). We thus need visual elements that support comparison, such as length. Additionally, bar charts which utilize length to display information are a familiar type of graphic adding to the readability of the data graphic.

4. *Highlight the key messages*: The data graphic was reviewed and the portions of the representation identified that most clearly illustrated the key messages. Arrows containing a summary of key messages were added pointing to the portions that most clearly illustrated the respective key message.

5. *Refining the visual representation*: To refine further the data graphic, the checklist in Appendix 5.1 was followed. First, the data graphic should be standalone. A descriptive title with keywords highlighted in bold was added. The axes were clearly labeled with "Transmissiblity" and "Severity." Both axes have labeled just two categories: "low" and "high." The category "medium" was omitted as it seems redundant to explicitly specify "medium" as being between "low" and "high," and omitting one category reduced the label density. Further, a legend explaining the two colors used in the bar charts was added. Finally, the two models compared ("cohorted" and "non-cohorted") were explained at the side of the representation under the legend. A single font was used for all text but some variation in size was used to reflect the relative importance of the conveyed information.

Second, we considered the technical requirements for the data graphic. The representation uses black or greyscale to show all information. Only the borders of the arrows are red to attract attention. However, if the data graphic is reproduced in greyscale, no information would be lost and the messages are still clearly discernible. Using greyscale also enables optimized used of the png file format and thus a small file size.

Third, the data graphic was "polished" to improve both aesthetic and analytical appeal. Care was taken to align the seven sub-graphics to support comparison of values. Additionally, the

10.5281/zenodo.15331321

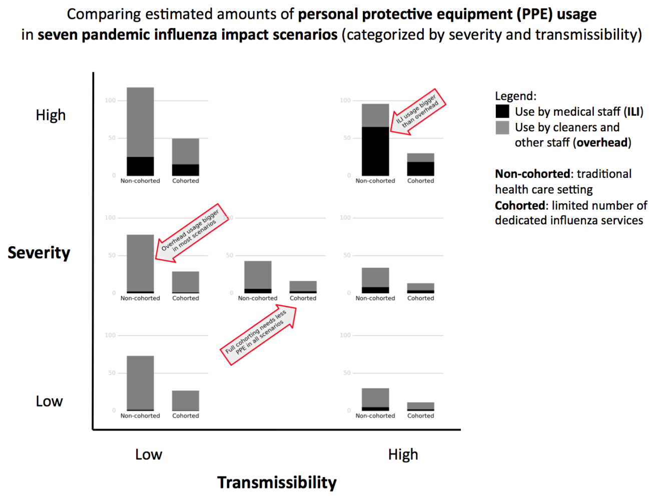


Figure 5.3: The new standalone representation communicating the key findings of the personal protective equipment stockpiling report [2].

grid lines of each bar chart was changed to light grey to reduce contrast. Grid lines support comparison, but dark lines, too many lines, or tick marks can be distracting (sometimes considered "chart junk" [4]). All bar charts use the same scales on the x and y axes further supporting comparison of values. The two grey values showing the two categories "ILI" and "Overhead" were chosen so that both are clearly visible, but also clearly distinct from each other.

5.1 *Summary*

Although only a relatively small pilot project, the feedback from both the scientists and the policy makers for the results of this project have been highly encouraging. The project has demonstrated how the application of simple data graphics design principles can substantially improve the coherence, clarity, and usefulness of data graphics in scientific reports to policy makers. Amongst the key findings of the project include the importance of:

- adhering to fundamental principles of data graphics design, including maximizing the data ink ratio and minimizing "chart junk," and selecting data graphics elements that enable rapid comparison of data along aligned scales.
- supporting multi-granular representations, that can be read both at an executive summary level by decision makers, while still allowing "drill down" low-level analysis and interpretation by policy advisers.
- facilitating display and sharing using a variety of technologies, including hard copy print-outs, emailing, low-resolution data projectors, and greyscale reproduction.
- interleaving of data graphics and narrative elements in the report, ensuring that key data graphics and report writing both highlight and reflect the key messages of the report.

The results also suggest a wide variety of key options for future work building on this successful pilot, including:

- Extension to wider domains: This pilot project has only considered one application domain, health. However, the problem addressed by this work—the effective communication of complex scientific information to policy makers, has wide applicability. Current work is already orienting towards extending these results to wider domains, including water resources and smart cities and urban planning. An ARC Discovery Project, currently under consideration for commencement in 2015, has been submitted to the ARC with this specific focus. The current MSoG Incubator Grant was instrumental in developing and submitting this major funding proposal.

- Development of infographics: The vital importance of narrative—"telling a story"—was an unexpected but clear outcome of the requirements analysis conducted in this project. Supporting the further development and integration of infographics with scientific reports, for example as the core of the executive summary of a report, warrants further investigation and work.

- Linking and layered reports: Amongst the other key outcomes of the requirements analysis in this project was highlighting the importance of multi-layered reports, which allow communication of high-level messages to decision makers, while still facilitating "drill down" analysis by policy advisers. Future work could profitably investigate technologies to assist with such multi-layered reporting, such as hyperlinked documents, linked data structures, or XML-based technologies that enable multiple different reports to be compiled from the same underlying documents.

- Fully interactive reports: Finally, moving beyond simple linked report, there is an opportunity to introduce even more interaction, for example even allowing policy advisers to execute and explore carefully constructed and restricted aspects of the full scientific model. As a stepping stone, the paradigm of "executable documents" might also be used to embed limited simulation and modeling capabilities within scientific reports. Such approaches hold the potential to allow policy makers to attain a deeper understanding of the results, sensitivity, and uncertainty associated with models, without overburdening reports with excessive data graphics and tables.

10.5281/zenodo.15331321

6 Bibliography

[1] J. McVernon and J. McCaw. Development of Options on how to Define the Concept of Pandemic Impact for Australian Purposes. Report on simulation modelling, For the Office of Health Protection, Australian Government Department of Health, 26th June 2012.

[2] L. Martich, K. Glass, G. Mercer, J. Ross, J. McCaw, and J. McVernon. Final Report on Estimated PPE Usage and Stockpiling Requirements. Report, For the Office of Health Protection, Australian Government Department of Health, 23rd Dec 2013.

[3] D. Borland and M. Russell. Rainbow color map (still) considered harmful. *IEEE Computer Graphics and Applications*, (March/April):14–17, 2007.

[4] E.R. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, 2nd edition, 2001.

[5] W.S. Cleveland and R. McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.

[6] A. De Raaf. Digital healthcare: Where are we now? `http://infographiclist.com/2011/10/01/digital-healthcare-where-are-we-now-infographic/`, 2011.

[7] University of Queensland Health Information Systems Knowledge Hub. Infographic: Health information systems. `http://www.uq.edu.au/hishub/infographic`, 2013.

[8] A. H Robinson, J. L Morrison, P. C Muehrcke, A. J Kimerling, and S. C Guptill. *Elements of Cartography*. Wiley and Sons, New York, 6th edition, 1995.

[9] Executive Healthcare Management (EHM). The role of infographics in healthcare communication. `http://h-t.com.au/the-role-of-infographics-in-healthcare-communications/`, 2013.

Checklist for representations

- □ Explanatory/summarizing title.
- □ Axes labels explanatory rather than symbols, abbreviations, or acronyms.
- □ Contains legend, legend entries explanatory.
- □ Font size legible, single font style throughout graphic, except bold/italic for highlighting.
- □ Representation can stand for itself, no additional explanations or captions needed.
- □ File size small (vector graphic or png), resolution maximum XGA 1024x768 (projector).
- □ Representation in black/white or greyscale, use of color sparingly as highlights or designed color scheme used, reproduction of color on greyscale printout works.
- □ Double encodings of the same information avoided (e.g., use of bar length and color to show the same information), except for instances of deliberate highlighting.
- □ "Chart junk" [4] removed, e.g., no avoidable lines, tick marks, tick mark labels, shadows, frames, background shading, decorations, ...
- □ Maximize data ink [4], i.e., of all the "ink" used for the representation (including frames, labels,...); most "ink" is used to show the data.
- □ Comparison between representations, parts of representations is encouraged, e.g., through aligned and equal axes.

10.5281/zenodo.15331321