Chapter 4: Design

The two studies in the previous chapter reveal opportunities for the use of information graphics in HTA. This chapter aims to show a selection of different information graphics that could be used to effectively present data to health policy decision-makers.

The transition from identifying a need to specifying a graphic that will address it is an intensely creative process. The first part of the chapter, 4.1, examines how information graphics work, and suggests a way of approaching the design of such visual presentations.

The second section, 4.2, provides 'specifications' for ten HTA information graphics. These give a quick overview of the information needs that they seek to address, a design sketch, and a brief description of the way in which the information graphic is intended to work. As time for developing such information graphics was limited, a panel of HTA experts was used to critique these specifications, and focus on the most promising among them.

The third section of this chapter, 4.3, provides a more detailed description of the five information graphics that were thought to be the most promising from their specifications. These have been designed 'iteratively', by showing rough drafts to researchers, and incorporating feedback into successive versions of the graphic. Real data from past reports has been used, so that the graphic is as close as possible to what would be used in practice.

These developed information graphics provide material for empirical research in the next two chapters, 5 and 6. These two chapters each test one of the developed graphics in detail, reported in a scientific manner.

4.1 Elements

4.1.1 Introduction

Information graphics often transform numerical data into visual elements (Macdonald-Ross & Waller 2000), such as the position of a mark or object (Bertin 1981) or the count of objects appearing (Neurath 1936). Some designers use multiple elements together to present information in a visual manner. The representation or transformation choices are often given to the viewer through a key, forming a temporarily agreed visual language between the designer and viewer. However, they can also rely on visual convention, such as using an arrow to suggest a direction of movement.

Few's term 'elements' (Few 2006) is used here for these methods of representing information visually, although other writers have used other terms such as visual 'grammar' (Ware 2004) or 'transformations' (Macdonald-Ross & Waller 2000; Neurath 1936). As mentioned in Chapter 1.1.2, several different design writers have suggested different lists of these elements, each providing their own palette of methods that can be used to represent information in visual form.

Bertin, who tends to focus on data maps, writes of x-position and y-position being two separate elements. He states that these two elements are always used, and provides six others which can show a third dimension of data (see Figure 4.1 – 1).

Ware has a more comprehensive list of elements, and dispenses with rules about which can be combined. He uses both categories and subcategories (Ware 2004):

Form

- Line orientation
- Line length
- Line width
- Line collinearity
- Size
- Curvature
- Spatial grouping
- Blur
- Added marks
- Numerosity

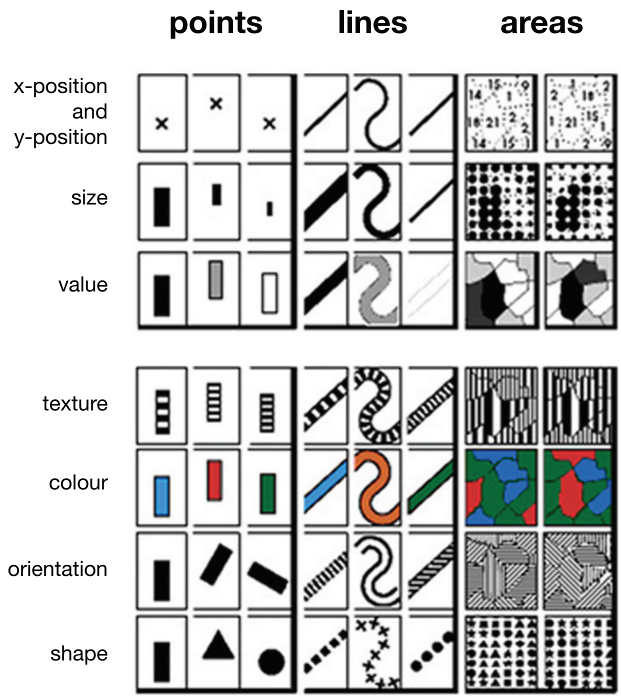


Figure 4.1 – 1
Bertin's list of elements (from Bertin 1981)

Color

Hue

Intensity

Motion

Flicker

Direction of motion

Spatial position

2D position

Stereoscopic depth

Convex/concave shape from shading

Stephen Few, who focusses on interactive information graphic 'dashboards', finds that he needs to adapt this list somewhat, reducing the number of categories, and removing the two-level heirarchy (see Figure 4.1 – 3).

These different designers have different lists of elements, which are useful for their particular information graphics – in Bertin's case, data maps, and in Few's case, information dashboards. These differences in the elements used might also be explained by the fact that designers have personal styles and ways of working, often producing different responses to the same brief (Brath 2008). It might be useful for the designers of information graphics to produce their own lists that fit their own personal ways of working and thinking, and perhaps also the context of their work.

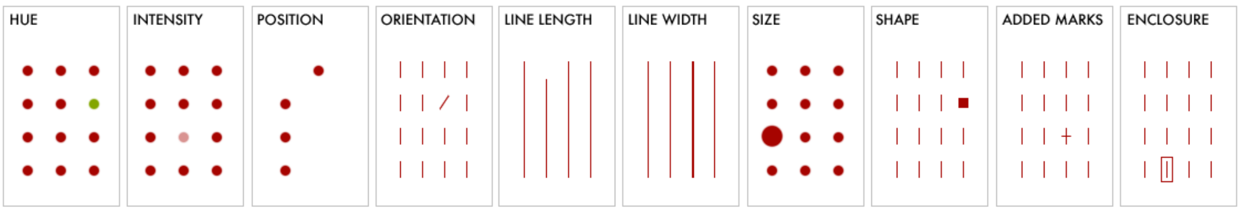


Figure 4.1 – 3
Few's list of elements (from Few 2006)

4.1.2 Personal listing strategy

For the design of HTA information graphics, a personal list of elements for the design of information graphics was constructed. Information graphics produced by other designers were collected, and an attempt was made to categorise the elements that they had used to present information.

While a systematic search for information graphics to include in this collection could be attempted, the TAR review detailed in chapter 3.1 revealed that the current use of information graphics in HTA is too limited to produce a comprehensive list of elements. Therefore the information graphics used to construct the element list were from a wide range of disciplines and areas. These were collected from numerous sources, including books, newspapers, journals, conferences, lectures and websites.

For example, the information graphic in Figure 4.1 – 4 (Senseable City Lab 2011) uses several graphical elements. These could be categorised in many different ways. Bertin might suggest that the graphic uses:

- **x-position** to show the longitudinal position of each area of New York
- **y-position** to show the latitudinal positions
- **colour** as his 'third dimension' to show which countries had been connected to

However, this categorisation does not account for the amount of calls to each country, which is shown on the graphic using the vertical sizing of the bars within each square.

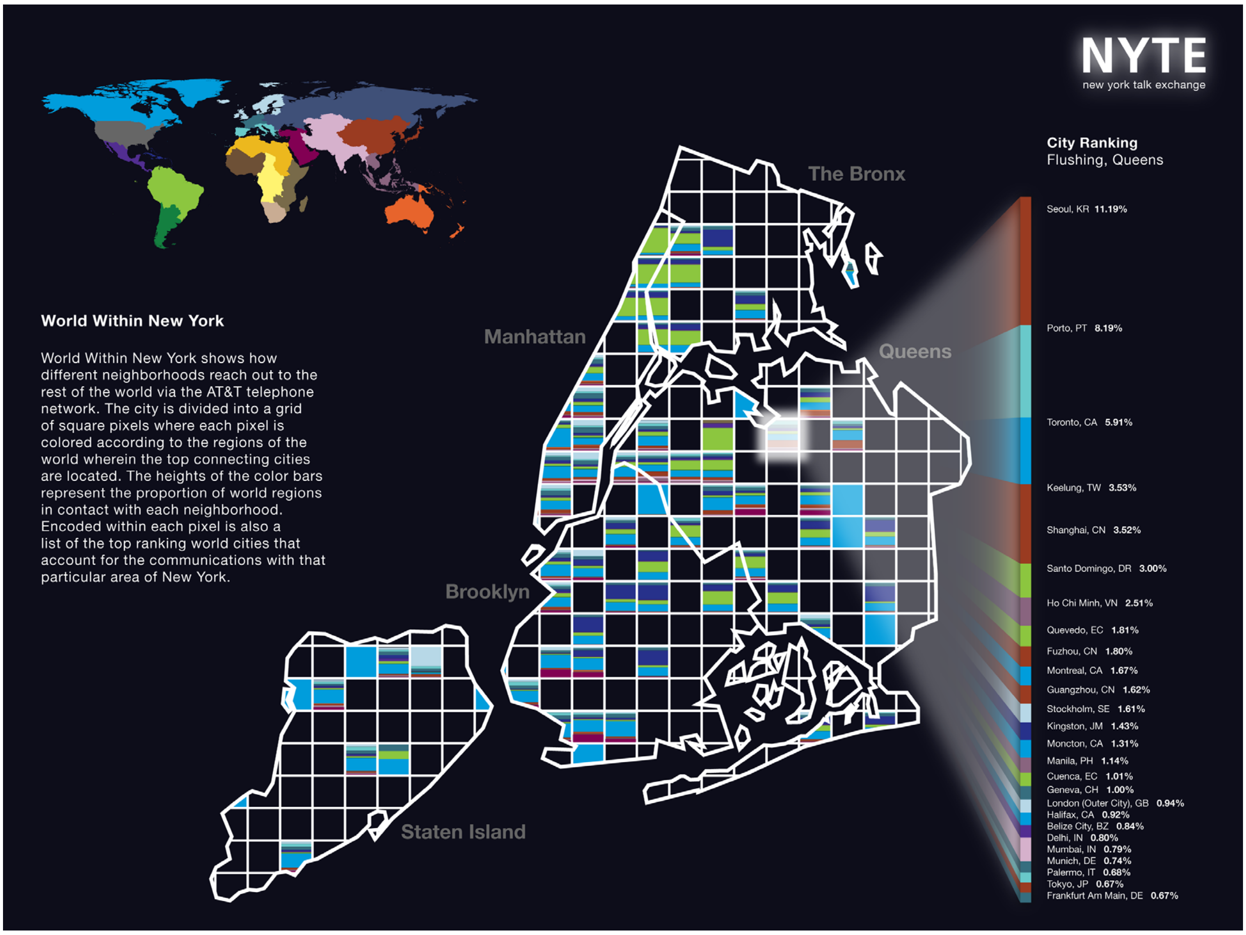


Figure 4.1 – 4
New York Talk Exchange – Senseable City Lab, MIT

Few might categorise the graphic as using:

- **position** for the areas of New York
- **hue** for the different countries connected to
- **size** for the amount of calls to each country
- **enclosure** for the currently selected square

From the list of elements used in this thesis, the New York Talk Exchange graphic might instead be categorised as using:

- **2d position** for the areas of New York
- **colour (for distinction)** for the different countries connected to
- **1d size** for the amount of calls to each country
- **highlighting** for the currently selected square

It should also perhaps be acknowledged that the designer of the New York Talk Exchange graphic has chosen to present textual and numerical data alongside the graphical elements. As these are not just for the purpose of providing a key to the visual elements, but actually form an integral part of the data presented, the graphic can be said to incorporate further elements:

- **numerals** to show the exact percentage figure for the calls
- **text** to show which cities are connected to in the selected square

The designer has also made use of some other typical graphic design elements to aid the user in navigating the visual space of the graphic. Therefore, the graphic could also be said to employ:

- **font weight** to visually separate the textual descriptions of the cities connected to from the percentage values of calls
- **grouping** as proximity is used to link the names of the different areas in New York to their corresponding locations

In the information graphic in Figure 4.1 – 5, the designer has included individual photographs to make a powerful and striking visual presentation. It shows each of the British soldiers killed in the second Iraq war from 2003 onwards. The graphic was categorised for the thesis as using:

- **count** for a number of objects appearing, in this case photographs
- **physical** representation for the use of photographs of the people being counted
- **grouping** as these photographs are grouped by year
- **numerals** as the total number of deaths each year is presented numerically

While text is used in the graphic, it is for the purpose of giving a scale (the years) and titles. Therefore it would not be included as an information graphic element in this case.

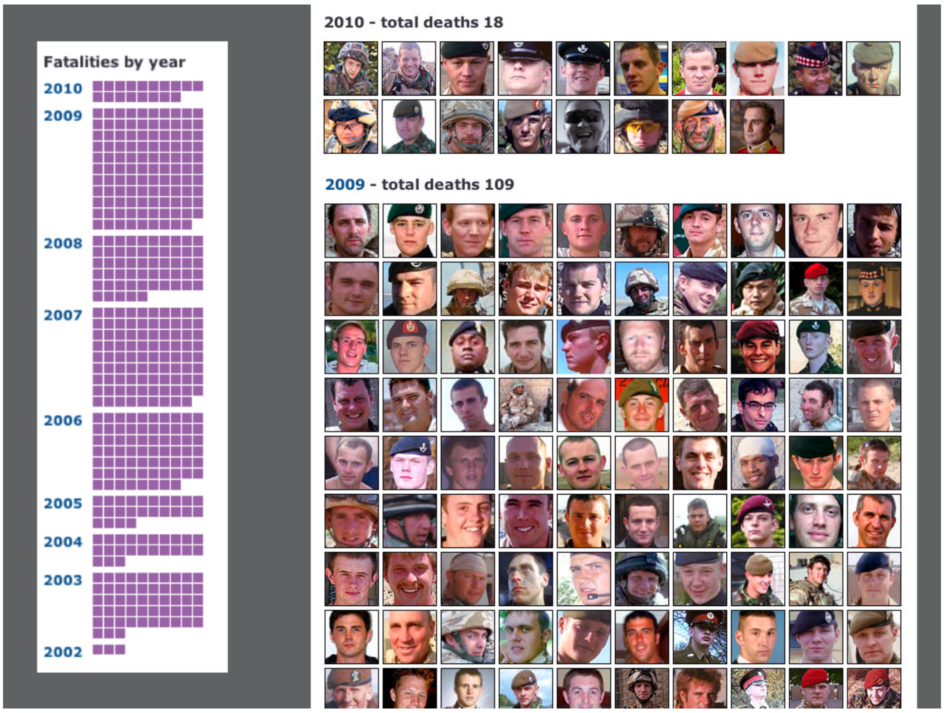


Figure 4.1 – 5
BBC news graphic – casualties in second Iraq war, by year

4.1.3 Elements

The following tables provide a list of elements that was used as a basis for designing the information graphics in Chapters 4.2 and 4.3. Table 4.1 – 1 gives a detailed list of static elements that could be used in a printed document. Table 4.1 – 2 shows some non-graphical elements that can be incorporated into information graphics. Table 4.1 – 3 gives a brief list of active elements that could be used for screen-based information displays, but these are not the main focus of the research here.

The list is considerably longer than those used by Bertin, Ware and Few. As already mentioned, this can be explained, at least in reference to Bertin and Few, by their more specific remits – data maps and information dashboards. In Ware's case, the smaller number of possible elements used could also be due to the limited range of information graphics and applications with which he is interested. However, for all three writers, it could also be that the lists stem from their individual theoretical discussions (and in Few's case, from Ware's theory on which it is based), rather than from the challenges faced by designers in a real-world context of practice. The list here is based instead on a taxonomic approach, looking at which techniques exist and are used by a very broad range of designers.

It is of interest to note that many of the elements mentioned in Bertin, Few and Ware's lists were found 'in the wild', and are therefore included in this new taxonomic list. However, there are some alterations. For example:

> Ware's 'line orientation' and 'line colinearity' are combined in a single element – orientation.
>
> Several elements that Ware applies only to 'lines' were used in the surveyed works on other objects, and therefore 'line length', 'line width' and 'size' are combined (and, indeed extended) in '1D size', '2D size' and '3D size'.
>
> 'Curvature' and 'added marks' are included within 'shape'.
>
> 'Spatial grouping' is included simply as 'grouping', since other techniques (such as enclosure) are often used.
>
> 'Blur' was not used in any examples found, so its value has not been sufficiently demonstrated for inclusion in this list.

This list has proved sufficient for creating some information graphics within the context of HTA, and it is hoped that it could be applied (and expanded) in other contexts of use.

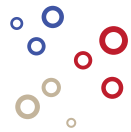
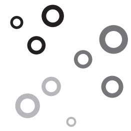
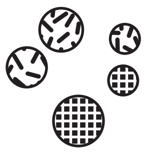
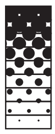
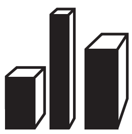
Element	Description	Example
Colour (distinction)	When colour is used for distinction only – to distinguish between sets, for example	
Colour (data)	When numerical values are represented with colour, on some kind of scale (graded or continuous) or colours are recorded	
Greyscale (distinction)	When shading is used for distinction only – to distinguish between sets, for example	
Greyscale (data)	When numerical values are represented with greyscale shading	
Pattern (distinction)	Use of hatched areas or dashed lines	
Pattern (data)	use of pattern to suggest scale (common in laser printed documents – known as 'rosing')	
Size 1D	One-dimensional sizing of objects (as in bar graphs)	
Size 2D	Two-dimensional sizing of objects (including bubble charts and pie charts, which rely on area)	
Size 3D	Three-dimensional sizing of objects. Can be clumsy or confusing to represent on paper or screen.	
Position 1D	Objects placed on a single axis. Data represented by position on a line, circle or other shape, from one point (often representing 0) to another. Can be: **nominal** (a list with no numerical significance, often alpabetical) **ordinal** (list from highest to lowest for example) **interval** (position represents numerical value) **ratio** (interval scale, must begin at a 0 value meaning nothing)	

Table 4.1 – 1
List of static elements produced for thesis

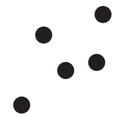
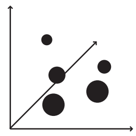
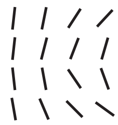
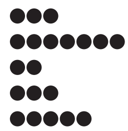
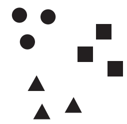
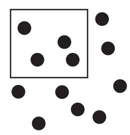
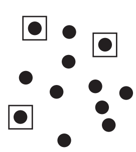
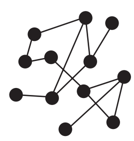
Element	Description	Example
Position 2D	Data represented using 2 axes, like a scatter plot, line graph or points on a map	
Position 3D	Data represented using 3 axes. Can be clumsy or confusing to represent on paper or screen	
Font weight	A graphic design technique that is applicable to information graphics. Numerical data can be represented by using a more or less bold weight of typefaces. This is particularly effective when used with a typeface such as Univers that has many different font weights	UltraLight Light Regular **Medium** **Bold**
Orientation	Use of rotated objects. Often used for showing flows on maps, particularly in meteorology	
Count	Where the number of objects that appears is significant, as in the Neurath's Isotype diagrams	
Shape	Representation of different things using shape. This might include the repeated use of logotypes or pictograms that must be learned	
Symbol	Representation using an assumed convention, such as arrows for direction	
Grouping	grouping elements, for example using an outline or enclosure	
Highlighting	The use of border etc to highlight element/elements	
Linking	Physical joining of elements using a line, etc.	

Table 4.1 – 1
List of static elements produced for thesis (continued)

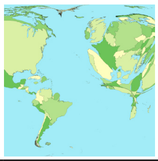
Element	Description	Example
Physical representation	photographs and illustrations	
Homunculus	proportional distortion	

Table 4.1 – 1
List of static elements produced for thesis (continued)

Element	Description
Numerals	Data is displayed in numbers, excluding scales and titles.
Numeral scale	Numbers are included to explain a scale.
Text	Graphic includes textual data (excluding titles and scales)
Text scale	Text appears on a scale, including axis labels

Table 4.1 – 2
List of non-graphical elements produced for thesis

Element	Description
Mouse position	mouse position affects the display
Object motion 1D	Data is represented by the motion of objects over time, in a single dimension
Object motion 2D	The motion of objects in two dimensions
Object motion 3D	The motion of objects in three dimensions (this can be clumsy or confusing to represent on a two-dimensional screen)
Motion attention	Calls attention using motion or quick resizing
Linear time	The use of linear time, such as a video playing at a constant speed
User time	The user can control the speed or direction of video playing, or skip to a certain part of the video
User parameters	The user can control what input data are used
User parameter selection	The user selects inputs from a list
User scale	The user can control the scale of the graphic
User sort	Data can be reordered by the user
User position 1D	The user can control the position of the data in a single dimension
User position 2D	The user can control the position of the data in two dimensions
User position 3D	The user can control the position of the data in three dimensions
User highlight	The user can select an element or elements, or the results of a query are highlighted
User search	The user can type text or numbers to search for them
User detail	The user controls the level of detail
User~	Many other static elements could be controlled by the user.
Hyperlink	A web link to other sites or different parts of a site
Sound	Not strictly a graphical quality, but can be used to present data or call attention to certain conditions being reached.

Table 4.1 – 2
List of active elements produced for thesis

Such lists can never be exhaustive, particularly as technological advances continually provide new ways of displaying data. However, they serve as a useful focus for a designer, helping them to consider the possibilities of visual information presentation from a personal perspective. The elements detailed here have been used to develop the new information graphics described in the rest of this chapter.

4.2 Specification

The next phase in the proposed design process for information graphics in HTA is called 'specification' (see Figure 4.2 – 1). If the 'information needs' phase maps roughly onto the 'discover' phase in the Design Council's generic design process model, the 'specification' phase maps roughly to their 'define' phase. The objective is to find a specific information need and document this, but also to explain a way of addressing this need with an information graphic. This process was performed 10 times, which resulted in 10 different specification documents, detailing a particular technique that could be used to address an information need. These documents provide a basis for discussions between the 'expert communicator' (the information graphic designer) and the 'content experts' (HTA researchers).

It should be noted that in this case, an individual designer (the PhD candidate) performed all the design work. If a larger design team, or a design agency external to the HTA team was used, this step could be broken down into further stages – such as producing a 'brief' that outlines the problems to be addressed, and more concept sketches and rough designs that might address this brief, before the final specification is produced. In this model of the design process in HTA, these activities are assumed to be incorporated in the 'initial design work' arrow, that links information needs with the specification. The next major node

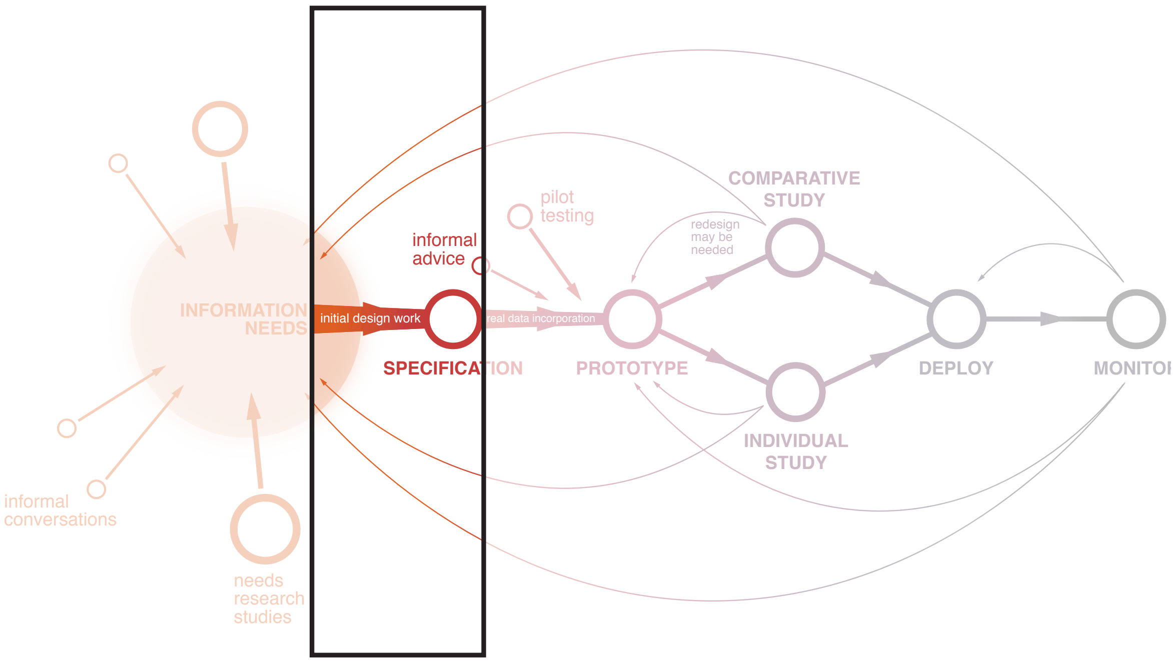


Figure 4.2 – 1
The phase of the proposed design process detailed in Chapter 4.2

on the diagram is this specification stage, as it is the next major interface point between the 'expert communicators' and the 'content experts'.

Each of the 10 specification documents produced at this stage in the PhD included a textual description of the problem addressed, a rough design sketch and textual description of the information graphic – including notes on different versions that could be produced, depending on media used (such as if colour or interactive versions were possible). A summary of each of these specified information graphics is given in this section. These show the origin of the problem(s) that they address, and provide initial ideas about existing data that could be used to make a prototype information graphic. Each specification then goes on to explain how the information graphic will present this data, showing the individual design elements, using rough design specification sketches.

Each of the specifications was subject to an informal critique by the supervisory team for the PhD. These discussions included the PhD candidate, who has training in graphic design, acting as the 'expert communicator' and three academic experts in the field of HTA, acting as 'context experts': a health economist, a mathematical modeller, and a review group director. In this situation, the three experts can be regarded a proxy for the 'client' that would usually be called upon to agree the design specification. One of the three HTA experts also sits on a NICE appraisal committee, so can offer a 'user' perspective, but all three have familiarity with the users and their needs through their work. Summaries of these informal critiques are presented after the individual design specifications.

4.2.1 Graphic 1 – Small multiple techniques including Sankey diagrams for overview of studies in a systematic review

4.2.1.1 Background

Systematic reviews in technology assessment reports often include many trials. In NICE appraisals that consider several subgroups or interventions, a review may use 50 or more trials to provide evidence of clinical effectiveness. This data is usually presented in very large tables, often combining text and numbers.

As the interviews presented in Chapter 3.2 suggested, decision-makers have limited time to look at the reports. A graphical overview which would enable the reader to understand such a large volume of information more quickly may be helpful for the larger systematic reviews in NICE technology appraisal reports.

4.2.1.2 Description

This information graphic uses several visual presentation techniques to reduce the page space needed to display the data. It then presents each of the studies in a subgroup in one row of a tabular display, enabling the viewer to see some information on each included study. In this case, the following information is included:

- Main author
- Date of publication
- Mean age and/or age range of participants
- Location of the trial
- Sample size and basic study design
- Length of follow-up
- Outcome measures used, and whether these were statistically significant

The country in which the study is performed is provided as a continent shaded on a globe, rather than a textual name. This is intended to be quicker to read.

The Sankey diagram shows the number of people in the trial, along with its design. The first three shown here are: randomised controlled trial, cohort study and pre/post study. The thickness of the lines is proportional to the trial's *N* value. The line following the diagram shows the trial's follow-up.

An overview is given of the ages of participants in each trial, giving the mean and range. Care must be taken that the graphical form of this part of the graphic does not look like a forest plot, to avoid confusion.

Outcome measures are presented with circles on vertical lines. This should allow individual outcomes of interest to be viewed by running the eye down a single line.

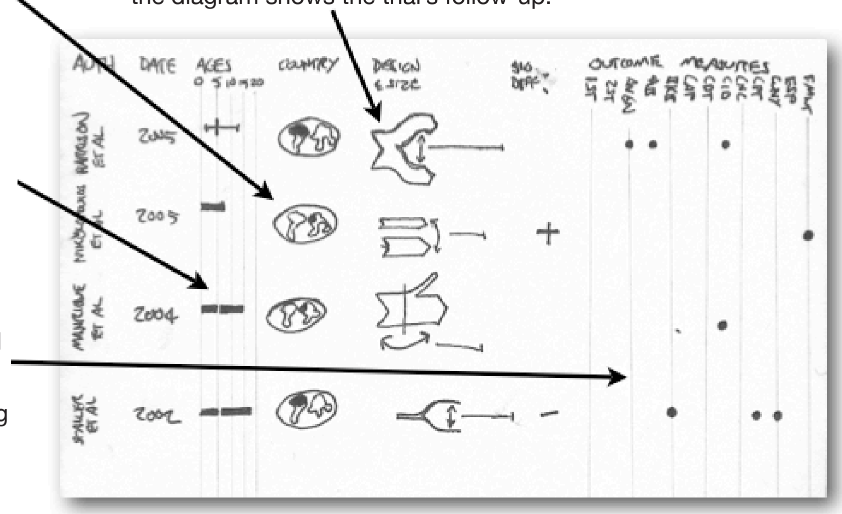


Figure 4.2 – 2

Small multiple techniques including Sankey diagrams for overview of studies in a systematic review

This graphical display should allow more data from more trials to be presented on one page, therefore enabling a more efficient overview of the information (see Figure 4.2 – 2). It combines Edward Tufte's ideas of using graphics to present more data in a small space (Tufte 2001) with Resnikoff's ideas of using the high bandwidth of human visual processing to enable faster assimilation of data into human cognitive awareness (Resnikoff 1989).

The graphic may not be able to include all the information on each trial, as there would not be space on the page for everything. While it is always difficult for a researcher to decide not to present certain information, it is also difficult for a decision-making committee to see the overall message in the data if too much detail is given. This graphic attempts to strike a reasonable balance between these two, by giving as much information as possible on key study characteristics, but preventing information overload.

4.2.1.3 Expert critique

The graphic was generally well-received by the three HTA experts performing the role of the 'clients' in the specification exercise. They felt that it gave a good overview of the information. One called it "A visual synopsis".

It was suggested that it would be useful for reviews that have a large number of outcome measures. These are often split up into categories in such cases.

One of the experts said of the graphic that: "There's so much useful information here, and you get it all in one place." In general, it was considered a useful way of getting a quick overview systematic review and outcome data.

4.2.2 Graphic 2 – Two-way sensitivity analysis matrix / bubble chart

4.2.2.1 Background

One very common practice in HTA that includes some kind of economic model is 'sensitivity analysis'. This comes in several forms, but one of the simplest ('one-way deterministic' sensitivity analysis) is to change a single input parameter's value in a model, such as altering the quality of life gained by using an intervention, to see the effects of this change on the results. This is important to consider, as almost all inputs to a model will be based on uncertain evidence.

However it is difficult to see, from a one-way sensitivity analysis like this, effects that could be compounded by alterations to other variables (correlations). Instead, it is possible to show the effects of altering two variables simultaneously with a simple graphical technique. This might be used to detect such relationships in a model, and/or to show this to the decision-makers.

4.2.2.2 Description

A matrix structure is used to show the effects of altering several different variables, two at a time. Each box in this matrix would contain a solid black circle, whose area related to the outcome of interest. In the case of Figure 4.2 – 3, this is the ICER value (incremental cost-effectiveness ratio).

One variable is presented horizontally, and the other vertically, forming the matrix of individual boxes. The 'base case' of both variables set to their initial values* are shown in the central box, with a medium-sized circle. Moving away from this central box either horizontally or vertically would represent a change in the corresponding variable. Moving diagonally would represent a change in both.

* Which are often obtained from systematic reviewing in the case of quality of life values or cost data.

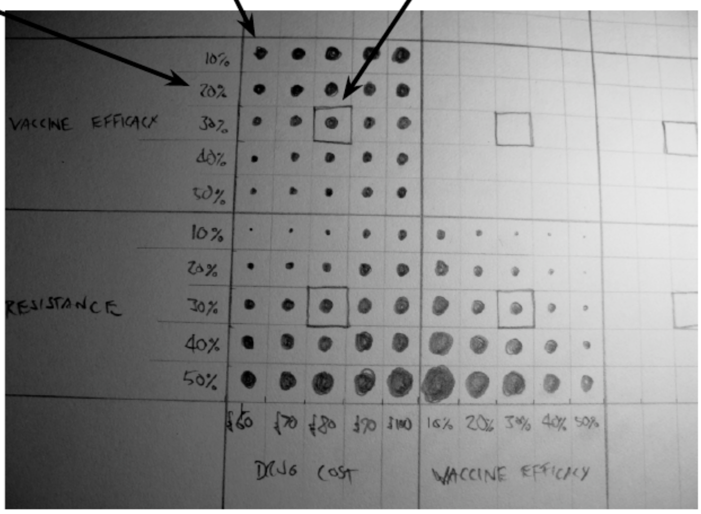


Figure 4.2 – 3
Two-way sensitivity analysis matrix / bubble chart

Using several matrices, arranged in a 'step' pattern, should allow all combinations of up to about five or six variables to be presented on one page. Seven or eight variables might be shown over a two page spread.

If one set of circles became very large or very small, this would indicate that the model was very sensitive to variations in that variable, and careful attention should be paid to how much evidence there was for this value. If a circle became very large or very small towards the corner of one of the matrices, it might show a possible correlation between the two variables used in that matrix. This information would not be shown in a one-way sensitivity analysis, and might be masked by the complexity of a probabilistic sensitivity analysis (see Chapter 4.2.3).

4.2.2.3 Expert critique

This graphic was discussed at quite some length, and proved controversial.

It was noted that the ICER might not be the most appropriate main output. Net benefit* would be another option. This would provide a potentially interesting role for interactivity, in that a slider could control the decision-makers' willingness-to-pay threshold value for this.

* The current value of the costs of adopting an intervention subtracted from the current value of the benefits.

It was not thought to be clear that this graphic would be an advance on what is already common practice in HTA. It was felt that a two-way sensitivity analysis is often not enough for a committee. Comparing three variables would be a more difficult challenge, and such three-way sensitivity analyses would be a more significant need. It was suggested that committee decisions often hinge on three variables that have both a large effect on the model and are uncertain (generally due to lack of evidence).

It was also considered that this graphic addresses a problem that is not specifically in the domain of HTA, but can be thought of as using a system to display multivariate data. In this sense, it's likely to be competing with work that's already been done.

It was acknowledged that currently the committee sees either everything or nothing. The probabilistic sensitivity analysis includes variation in all variables, and the one-way sensitivity analysis shows individual variables. However, there are other techniques for presenting two-way sensitivity analysis already, so this graphic was not thought to be worth pursuing further. A way of displaying a three- or four-way sensitivity analysis might be of more use, however.

4.2.3 Graphic 3 – Parallel coordinates for probabilistic sensitivity analysis

4.2.3.1 Background

One of the most sophisticated ways to address the uncertainty in a model is 'probabilistic sensitivity analysis' (PSA). This technique assigns probability distributions to all of the (uncertain) inputs to a model, rather than individual values as in a one-way deterministic sensitivity analysis. Many simulations are then produced, each time picking a different value from the distributions for each variable. While this can account for variables that influence each other when varied, the information needs interviews in Chapter 3.2 revealed that it can be challenging to incorporate the outputs of such a sophisticated method into a decision.

In the interviews in Chapter 3.2, four of the five interviewees mentioned that the current display of PSA can be confusing for (at least some) decision-makers. One of the main concerns was in incorporating the information from the PSA in the eventual decision. It was thought that, while it might be understood that an intervention could be said to have a certain probability of being effective at given willingness to pay thresholds, this could lead to difficulties in determining how this should affect a final decision. For example, if an intervention was 50% likely to provide cost-effective benefits at a threshold of £30,000 per QALY gained, it would be challenging for many people to decide whether it should be approved or not.

The aim of this information graphic is to begin to show which individual simulations in a probabilistic sensitivity analysis were above or below a given ICER (incremental cost-effectiveness ratio) threshold. The idea is to show not just how many simulations were cost-effective, but to suggest which ones were, and lead to a knowledge of why this was so. This would allow analysts and decision-makers to see any regularly-occurring patterns which might suggest specific conditions under which an intervention might be approved or rejected. Presenting this information visually might allow a committee to quickly see conditions which might arise in the future (perhaps in light of new evidence) in which their decision could be usefully reconsidered.

4.2.3.2 Description

This information graphic uses the 'parallel coordinates' display, which are usually interactive techniques. They are used to show many things, or instances of a thing, that have multivariate characteristics (Inselberg, 2009). The central premise is that several vertical bars are shown, each of which represents a characteristic. A set of 'cases' are shown as lines running accross the diagram, which represent individuals with these characteristics in varying amounts. These cases touch the vertical bars at the appropriate point for their value of each characteristic.

Usually, parallel coordinate displays are presented on screen, or using projected interactive media. This means that each characteristic can be used as a filter, with the user defining minimum and/or maximum levels. They are frequently connected to maps or other displays in a 'multiple coordinated views' format (Maciel et al. 2008). To be included in a printed or PDF format HTA report, an interactive tool like this could be used by an analyst, and any interesting observations could be shown as a still 'screenshot'.

The information graphic presented here is a set of parallel coordinates, showing each simulation in a PSA (see Figure 4.2 – 4). The vertical bars in this graphic correspond to a selection of uncertain variables in the PSA. The

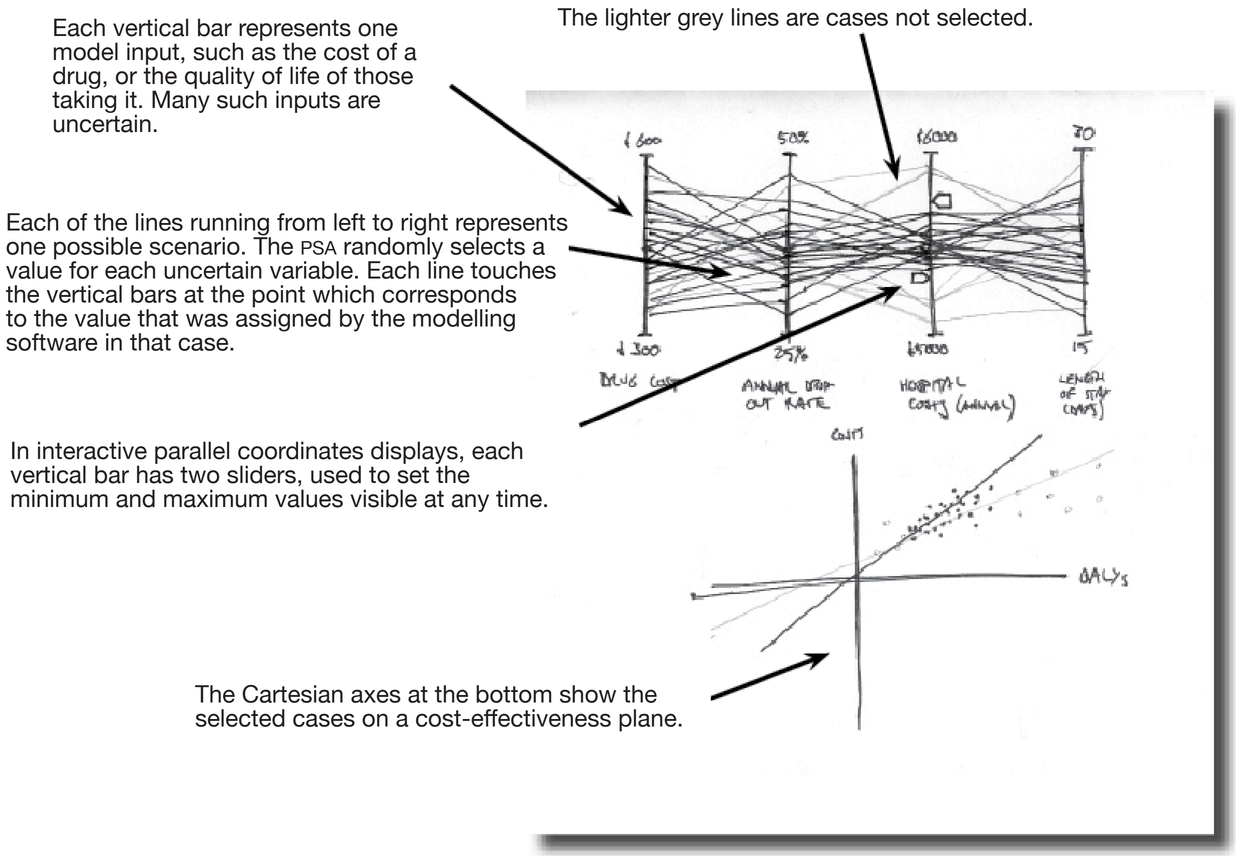


Figure 4.2 – 4
Parallel coordinates for probabilistic sensitivity analysis

analyst would filter the results of the PSA, with sliders which sit on each bar. These would be used to set limits for the maximum and minimum values of each variable to be displayed. The simulations that lie within the limits set by the analyst would be shown on a cost-effectiveness plane included with the parallel coordinates display. The values lying outside the set limits could also be displayed, perhaps in a different colour or a light shade of grey, showing the effect of the limits against the whole group.

This kind of display would allow an analyst to present a graphical representation of each of the simulations that came out of the PSA, and show the extent of the influence of certain variables, or combinations of variables, on the analysis.

4.2.3.3 Expert critique

This information graphic required more verbal explanation than most of the others. It seemed that the specification document that had been provided before the panel session was not sufficient for this graphic. It left questions such as how the outputs were presented (using the cost-effectiveness plane below the parallel coordinates display).

One of the experts mentioned that they often found themselves wanting to see what happens in a PSA when 'taking out' one variable entirely (ie, by setting it to a single, certain, value). This would require recalculation of the entire PSA, which would probably be too computationally intensive to be performed live. It was also suggested that EVPI (expected value of perfect information) analysis might be more helpful than this recalculation approach (Hubbard 2010).

It was felt that it might even be 'dangerous' to be able to set display limits for different variables. It might give the impression that the variability is being reduced, instead of just filtering the already calculated results. It was expressed by one expert as: "you're not getting rid of any model behaviour that depends on correlation that's occurring between other variables, and less limited values of that variable." While one of the experts thought that using sliders to reduce the variability of a variable (with recalculation) would be useful, another thought that this would be methodologically flawed.

It was thought that the cost-effectiveness plane would be useful, as it would be possible to see where the cloud of points moved to. It had to be explained again that the graphical technique as specified wouldn't in fact shift any of the points, but remove (or fade) them.

It was noted that the parallel coordinates display might be a useful tool for the modeller to validate a PSA. If, for example, a normal distribution had been assigned to a variable, the parallel coordinates would quickly show whether the simulated values of this variable had clustered properly in the middle or not.

One of the experts suggested that it relied too much on interactivity, and that the static 'screenshot' version that would have to be printed in the report would not be as valuable as some of the other graphics.

Scalability (Thomas & Cook 2005) was mentioned. It was asked whether the parallel coordinates display would really be able to cope with 1000 simulations, This number should certainly possible, although the lines might need to be partially transparent to make sure that some of the most densely clustered variables are still understandable. Also, it was noted that 100 variables in the PSA (ie 100 bars) would not be practical, so some selectivity might be necessary in terms of what was presented.

One suggestion was that one of the sliders could be used to show the ICER rather than one of the input parameters. This would allow a user to select a certain ICER range, and show which of the highlighted cases within this range were noticeably more tightly clustered on the different input variables. Such groupings would give an indication that these variables were influential on the ICER. While it might then be possible to dispense with the cost-effectiveness plane part of the graphic, it was thought that this display might still be quite useful, as it would be familiar to many decision-makers.

After a somewhat lengthy discussion, the value of this information graphic was still unclear. It raises a few methodological questions, and its final form was not completely fixed, in terms of what should be displayed. The graphic was taken to initial development stage to see if it would be readable with many simulations, and more input variables.

4.2.4 Graphic 4 – Technology assessment report graphical overview

4.2.4.1 Background

It is not unusual for complex situations to arise in HTA work, in which several different parts of the reports have a large influence on the decision. For example, in the NICE multiple technology appraisal of Adefovir, Dipivoxil and pegylated interferon alfa-2a for the treatment of chronic hepatitis B (Jones et al. 2009), the report concluded that different treatment sequences were appropriate for different subgroups. However, the published clinical effectiveness evidence available did not relate to sequential treatment. In this case, it might have been useful to consider the quality and appropriateness of the evidence of clinical effectiveness alongside the conclusions of the economic modelling.

While executive summaries in TAR reports go some way to providing an overview of the report, they must focus on giving a quick overview of the most important information. Therefore, space may not be available to present very much of the actual numerical data from the report.

4.2.4.2 Description

A standard way of presenting large volumes of information for quick summary is the poster format, often used at academic conferences. The aim here is to use a poster to show an overview of the entire report (see Figure 4.2 – 5). This poster would give graphical presentations of the numerical information from the report, supported by carefully placed text, using information design techniques such as position, hierarchy, and linking, to show relationships in the data. This would include information from the background of the condition, the clinical effectiveness, and the different economic analyses presented in the TAR, but displayed in one large layout. In the case of the hepatitis B multiple technology appraisal, this might have made the reasons for the sequencing problem apparent more quickly than reading the entire document.

Previous research has shown that graphical presentation formats can enable the viewer to understand numerical information more quickly, particularly where the information is complex (Remus 1987). Our own research, presented in chapter 3.2, suggests that decision-makers in the UK HTA system have limited time to read TAR reports, so enabling faster information absorption might have inherent value. However, it is also possible that some decision-makers may prefer a graphical presentation format. Accepting the conclusions of the research in Appendix A, that people perform better with presentations that they prefer using, this format might then lead ultimately to greater understanding of information.

One of the chief disadvantages of this graphical overview is finding a way of presenting it to the decision-makers. While such posters could be printed

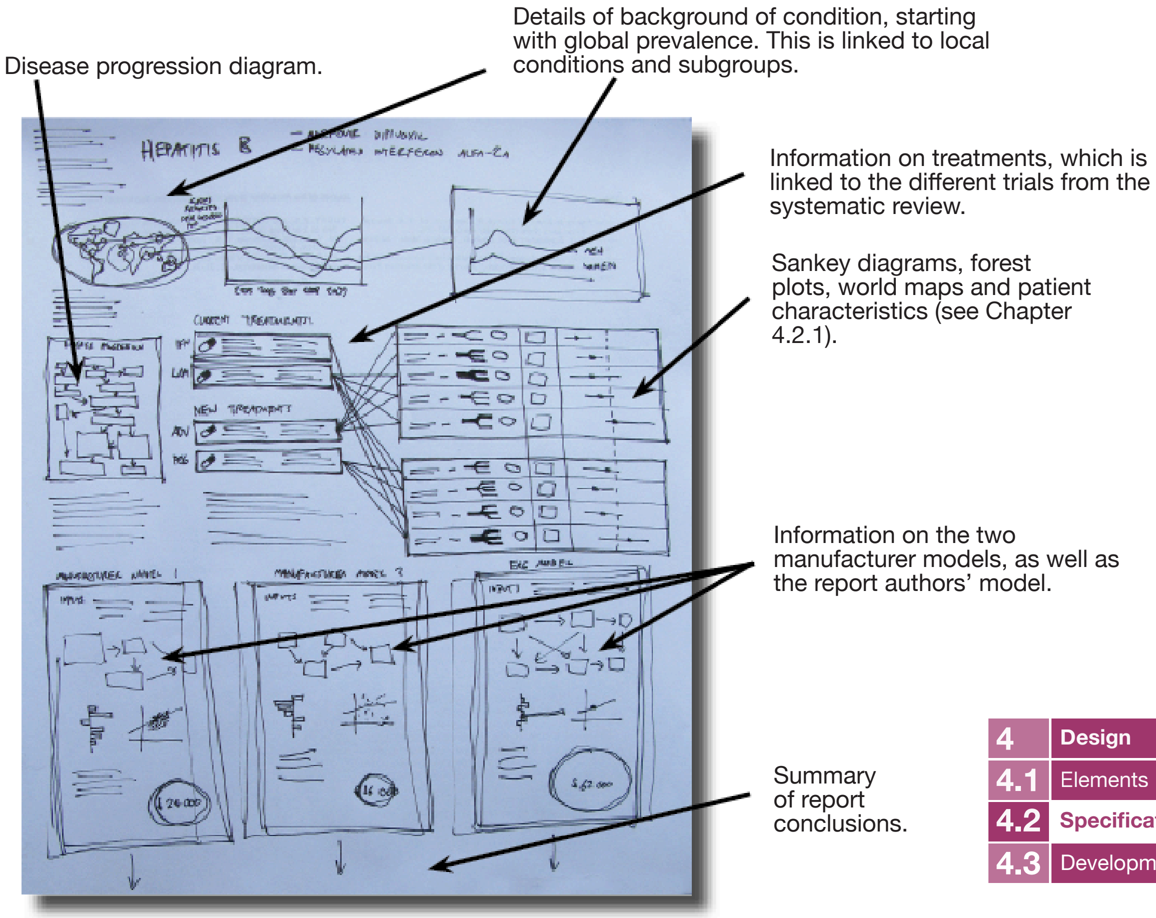


Figure 4.2 – 5
Technology assessment report graphical overview

on paper, displaying them in the venue for the appraisal committee would be of limited use, as the decision-makers are only likely to be present for the meeting itself. Sending such printed posters to each decision-maker would be expensive, and difficult for them to look at unless they could mount them on a wall or board. It would certainly not be possible to read on a train, as a report is. A poster could be sent electronically and explored by zooming and panning, for example using the widely available PDF file format, or websites such as (Prezi Inc. 2010). However, this would somewhat compromise the main strength of the graphic: being able to scan the data from the whole report by just moving the eyes over the poster. The poster could also be presented using the projector at an appraisal committee meeting, but would again require zooming and panning to have sufficient resolution to see the data.

4.2.4.3 Expert critique

Initially, this graphic was thought to have "no merit in the current format whatsoever." This was largely due to the difficulties inherent in distributing and viewing the design. It was observed that it would be necessarily less detailed than a several-hundred page report. Concern was expressed that, therefore, it wouldn't be able to replace anything.

One suggestion was that it could it be thought of as a kind of visual executive summary. However, it was noted that there is already a textual executive summary, which is quite valuable in its current form. It was suggested that the current executive summary has very little, if any numerical data in it, but this might only be true of badly written ones. While textual executive summaries may not include tables of numbers, the function of it is to give a view about what the key things are in the report, so selected figures should be presented within the text.

It was acknowledged that the poster would look visually attractive. However, caution was suggested, since a very aesthetically pleasing graphic that is not actually very useful could be unproductive, or even distract the viewer from the information contained.

One thing that is often unclear in HTAs is the degree of overlap in the sources of the cost-effectiveness analyses, and the effectiveness data. This poster display could provide that. For example, if the clinical effectiveness estimates were based on three trials, and the cost-effectiveness data came from just one of them, it might be possible to quickly get a visual sense of how these related to each other.

Overall, concerns still remained about how large a poster would be needed to display all of the data collected in an HTA. It might be an interesting task to attempt, particularly if the individual elements are being developed anyway, but it was thought that it probably wouldn't lead to anything substantial enough to be included in this PhD. It might be possible to develop a display like this if the individual components were to become very popular after some time. It was generally thought to be too complex and requiring too much change in terms of presentation media to be practical for this project.

4.2.5 Graphic 5 – Sankey Markov overview

4.2.5.1 Background

One issue, mentioned a number of times during the interviews in Chapter 3.2, was that decision-makers needed some way of seeing where costs and quality of life scores in a model came from, rather than just taking the ICER as a given result of the model structure and input parameters. This can be particularly problematic in the single technology assessment process at NICE, when manufacturers resubmit their report after a committee has issued its preliminary decision. Time can then be quite limited for decision-makers to get to grips with a new model.

4.2.5.2 Description

This information graphic is designed to show each cycle of the model in detail, with the costs incurred and QALYs lost in each one (see Figure 4.2 – 6). Situations in which cost savings are produced because of death occurring

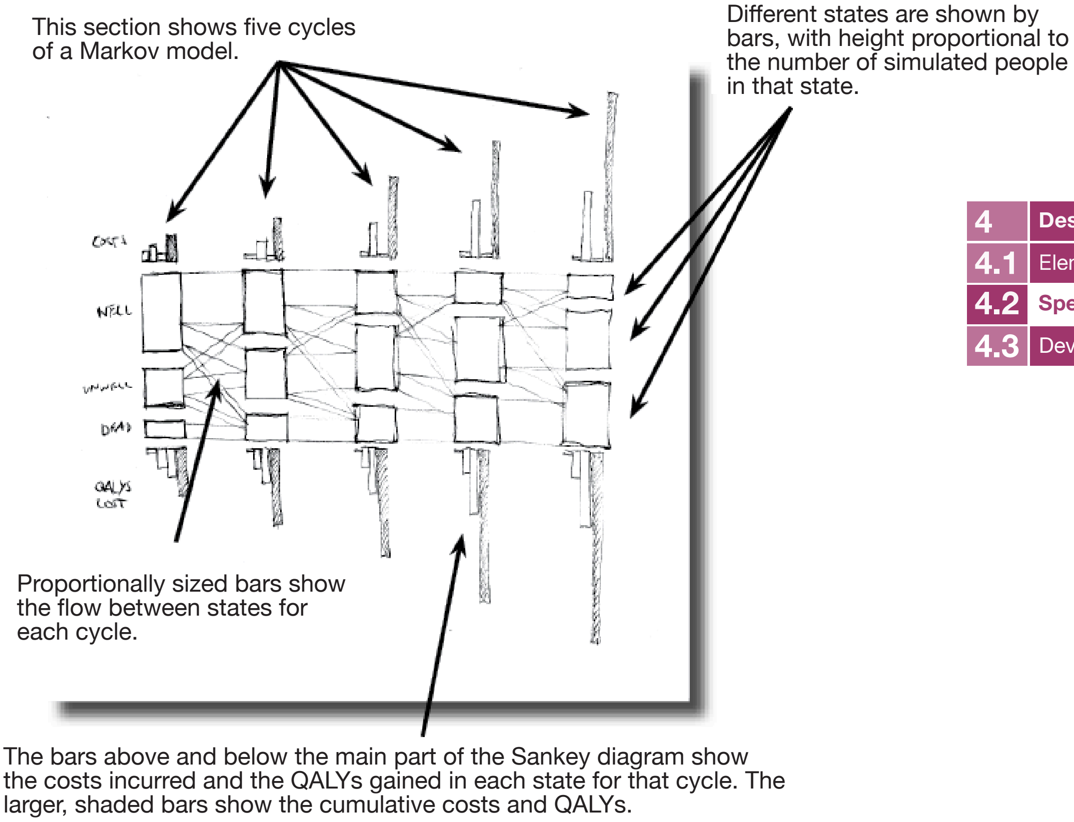


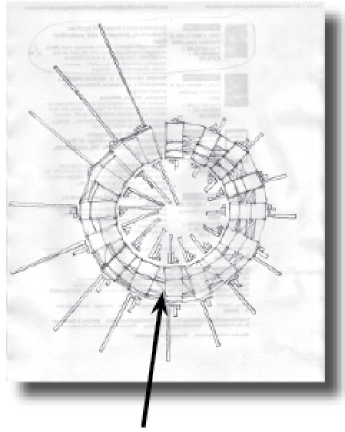
Figure 4.2 – 6
Sankey Markov overview

earlier in the model, or large costs incurred because of proportionally large groups remaining in costly health states should become more obvious than when looking at an aggregated metric such as an ICER. This should lead to a greater understanding of the temporal dimension of a model, and also allow a quick overview of the model dynamics.

4.2.5.3 Expert critique

The graphic was thought to address an important issue. An example was given of the screening of Hepatitis C (Stein et al. 2002), in which the different factors that contributed to the ICER were very important. This was also the case for the Rituximab model (Knight et al. 2004). Disaggregating the ICER is something that is often wanted in many appraisals, and should arguably be provided always.

Concern was expressed over whether the display would be practical, given the fact that many models run hundreds of cycles. Some form of clustering, such as grouping 5 or 10 cycles together, would be necessary to display what was happening on a single page of a report. It might be possible to produce a circular display, as shown in Figure 4.2 – 7. However, this would make it more difficult to compare the costs and QALYs produced from cycles in different parts of the model.



A model with many cycles could become very wide. A circular format might be more appropriate in this case.

Figure 4.2 – 7
Circular format

There was some discussion of whether this graphic would be an advance on current practice. This would generally be to include a state occupancy diagram, a comparative Cartesian time series graph, which shows how many people are in the different model states over time (see Figure 4.2 – 8). Reports often also include a separate description of how people flow from state to state, by presenting numerical counts of how many simulated people made transitions from one state to another. A separate Cartesian line graph might be used to show cumulative cost, and greyscale shading could be included on the state transition diagram to show quality of life, with black being equal to nearly dead. The benefits of the new information graphic would be that all this information would be brought together in one diagram. This would allow the accumulation of cost and QALY values to be compared more easily, and show more detail on how QALYs accumulate over time.

It was generally agreed that showing the source of costs and QALYs is a critical part of both the analyst seeing what their model is doing, and a committee understanding "where the ICER value comes from".

It was suggested that the graphic might be a good way of displaying information that was needed in the Hepatitis C screening report. In this HTA the benefits coming from the model were partitioned. Two different kinds of benefit were

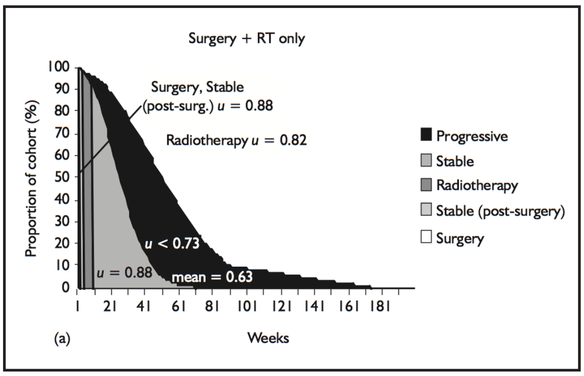


Figure 4.2 – 8
'State occupancy chart' from Garside 2007

assigned, one being small quality of life gains throughout the model, and the other being avoiding severe terminal consequences at later stages. The greatest benefit was shown to be in the small incremental quality of life gains in this case, and this kind of graphic might show the impact of such small, long-term gains quite easily. This could be particularly valuable, as survival estimates often “come from different places to utility estimates” – in that trials often measure one or the other, but do not report both. An example was given of the *Rituximab (MabThera) for aggressive non-Hodgkin's lymphoma* review (Knight et al., 2004)

There may be limitations with this information graphic, as models with many states may be difficult to display. Also, situations where progression is not mainly in one direction through the same series of model states might be more problematic.

Overall, the graphic was thought to need a substantial amount of work, but it was acknowledged that it dealt with such a key area that a developed version of this might be very useful.

This graphic was chosen for further development, as it “Gets more into the heart of the model” than any other way of displaying this information that is currently used. A small time investment might be necessary to learn how to use the graphic, but once decision-makers became familiar with it, it could provide a quick and effective way of displaying information about the accumulation of costs and quality of life over time for simulated people in a Markov model.

4.2.6 Graphic 6 – ‘Whirlpool’ display for enhancing tornado diagram in deterministic sensitivity analysis

4.2.6.1 Background

A ‘Tornado diagram’ is often used in technology assessment reports to show the results of one-way sensitivity analysis of an economic model. This vertically-arranged bar chart shows the effect on the ICER of different one-way deterministic sensitivity adjustments to the model (see Figure 4.2 – 8). It accepts negative values, as changes may increase or decrease the ICER value.

As the ICER is a ratio rather than an absolute value, these diagrams do not show the effect that these sensitivity adjustments have on the total costs and QALYS

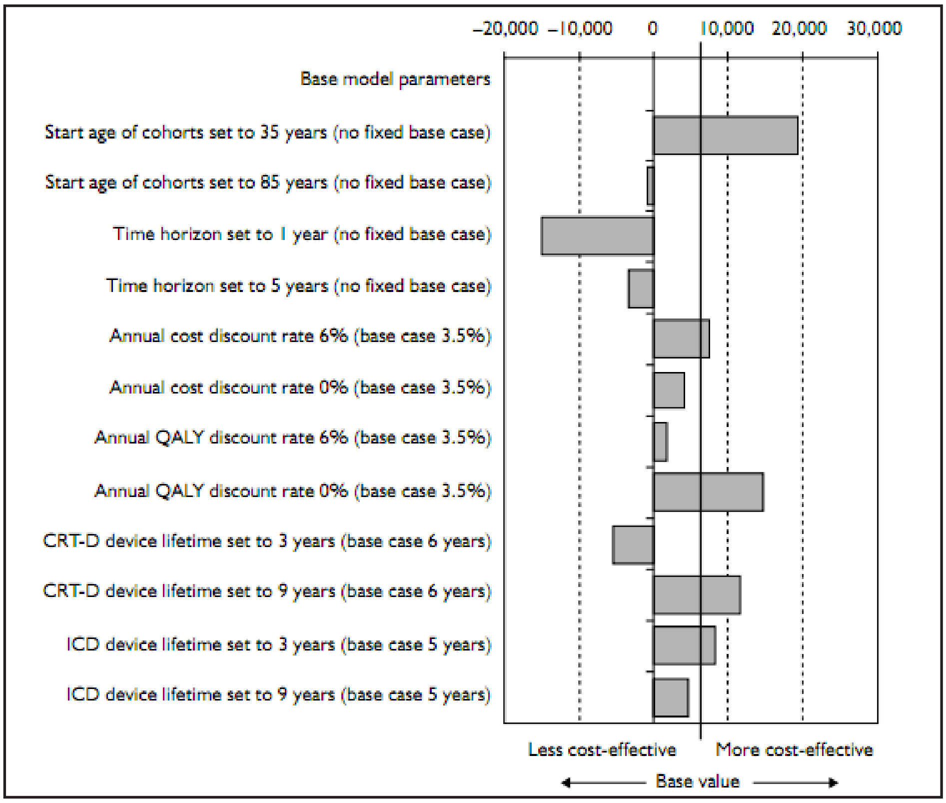


Figure 4.2 – 9
'Tornado diagram' showing the effects on the ICER of altering individual model inputs

accumulated under different interventions. These individual values might change without altering the ICER, where both costs and QALYs change by a proportionally similar amount. This information could be relevant to an HTA decision, but would not be shown in a standard tornado diagram. However, the ICER metric is still highly relevant, and should take (visual) priority.

4.2.6.2 Description

This new information graphic still shows the ICER change (positive or negative) with the usual 'tornado' bars. However, the 'whirlpool' display in the centre also shows the (positive or negative) change in the absolute costs and QALYs accumulated in the model. The distance from the mid-point ring shows how large a change in one of these values is, meaning that a large positive or negative change would appear as a dark area on the page, attracting the eye (see Figure 4.2 – 10). There are four 'quadrants' in each 'whirlpool', which represent the costs and QALYs for two different health interventions being compared.

4.2.6.3 Expert critique

The panel initially thought that showing the individual cost and QALY change for each intervention might not be necessary – showing the difference between

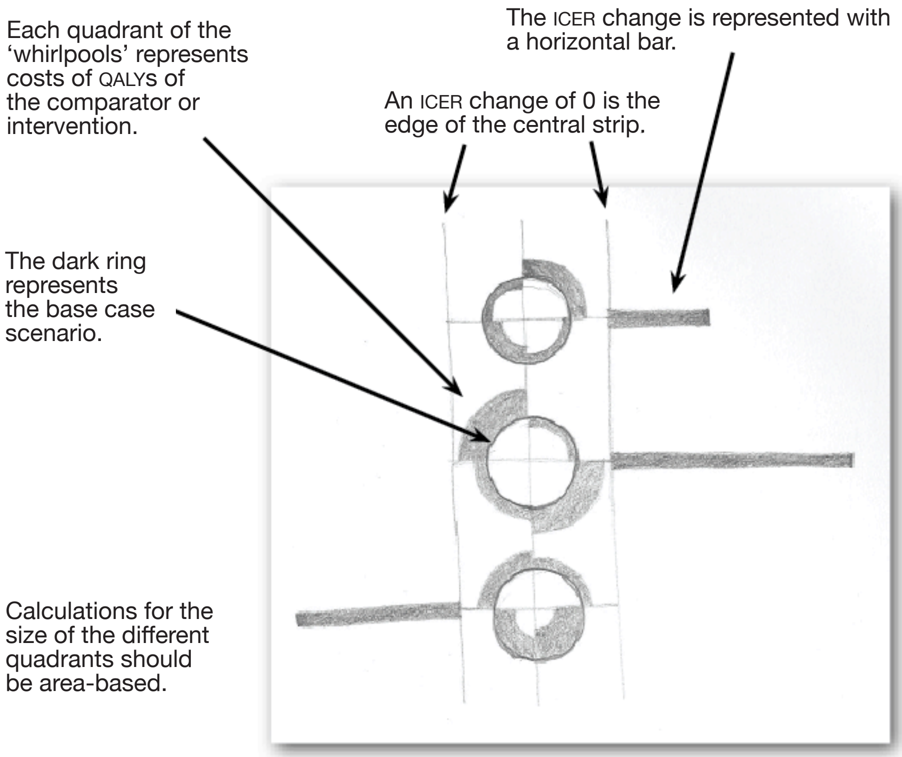


Figure 4.2 – 10
'Whirlpool' display for enhancing tornado diagram in deterministic sensitivity analysis

the two could be enough (the incremental cost or QALY change). However, this would not show equal changes in both comparator and intervention.

It was suggested that it would be more appropriate to show the changes on a Cartesian cost-effectiveness plane. Displays on a Cartesian plane are sometimes produced, which show the costs and effectiveness of two interventions, and the same values after a sensitivity adjustment is made (see Figure 4.2 – 11). However, it would be more difficult to compare the impact of the different sensitivity adjustments, as each would have to be displayed on a separate set of axes. Also, occlusion is likely to be a problem for very small changes.

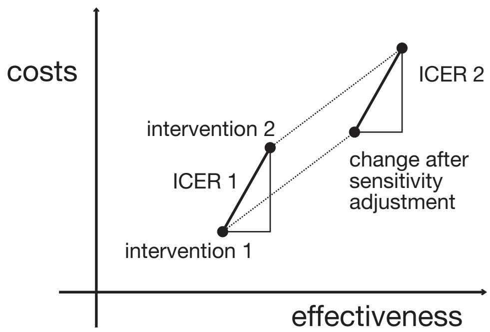


Figure 4.2 – 11
Movement of ICER (represented as the gradient of the diagonal lines) on a cost-effectiveness plane in response to a sensitivity adjustment.

It was generally thought that, while a better way of showing the outputs of one-way sensitivity analysis is needed, this is perhaps not the best way of doing so. This graphic was not chosen to proceed to initial development.

4.2.7 Graphic 7 – Survival synthesis bubble chart

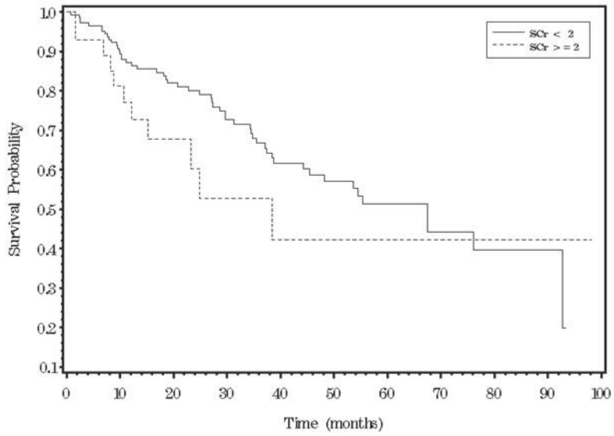


Figure 4.2 – 12
A 'Kaplan-Meier curve' showing the survival of two groups of people

4.2.7.1 Background

For many health interventions, the primary outcome of interest is the survival of the people receiving treatment. This is often presented in individual medical trials with a visual technique called a Kaplan-Meier curve (Kaplan & Meier 1958). Figure 4.2 – 12 shows a Kaplan-Meier curve, the two lines showing the overall survival of participants in the two groups of an RCT. In a systematic review, with many trials to consider, the survival curves provided by each report are difficult to compare side-by-side. Overlaying them on one set of axes, however, would be very confusing, particularly when one considers that even two curves on a Cartesian axes may be visually misleading (Cleveland & McGill 1984).

4.2.7.2 Description

For this information graphic, a bubble chart is used, which presents the survival data at five points, one for each year of follow-up in the trials. The number of surviving patients at any point is represented by the area of a circle (see Figure 4.2 – 13). This relies on the viewer's ability to judge relative areas of circles to impart the information. Luckily, the human eye is well trained to judge relative sizes. Over many years, we develop the ability to judge the distance of objects or which is the larger of two portions of food (Merleau-Ponty 1962). While this ability may be influenced by different shapes or perspective effects, it can be used by an information designer to display complex information, such as survival data from a large systematic review.

In theory, the use of area-based circles instead of the more usual curves allows for comparisons to be made easily between different trials (in a horizontal direction) and between years in one trial (in a vertical direction).

4.2.7.4 Expert critique

It was noted that the display in its current form does not handle censoring (where a person is lost to follow-up in the trial, and may be alive or dead).

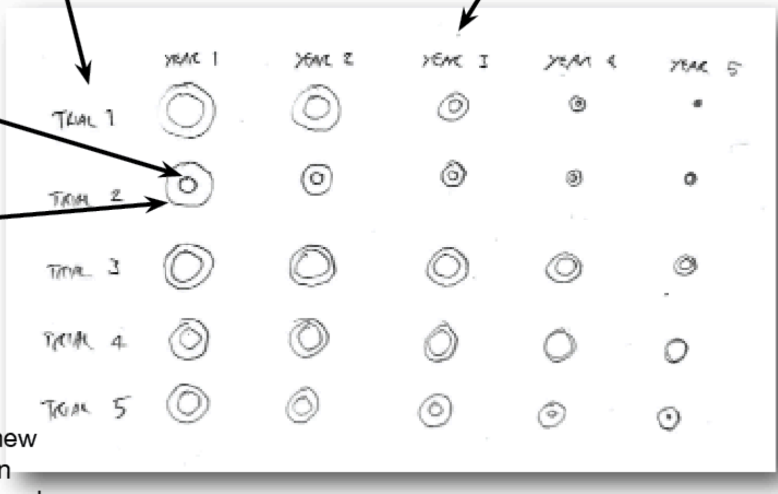


Figure 4.2 – 13
Survival synthesis bubble chart

This is normally shown in a Kaplan-Meier curve by adding a small mark to the curve when the patient is censored. The patient is removed from the group of alive patients at that time point, appearing visually to be counted as a death. Sometimes, multiple lines are plotted, one counting censored people as alive, and another counting them as dead, giving two diverging lines that indicate the minimum and maximum possible number of people alive. The true line is then expected to lie somewhere in the middle. The new graphical display, with its area-based circles, could use a similar technique, with the 'number possibly alive' data point indicated with a larger ring in 50% opacity outside the larger group, and a smaller ring in the same shade between the larger group and the smaller group, indicating the uncertainty caused by loss of follow-up. This uncertainty could also be presented with a linear gradient, from solid colour at the 'minimum possibly alive' point to white at the 'maximum possibly alive' point.

There was some contention in the group about how useful the display would be. One of the three experts liked the display, another preferred to look at multiple line graphs, and the other was unsure which they preferred.

This third person did like the way that the display allowed different trials to be compared, but was not convinced that judging a circle's area would be as accurate as judging the position of a line on a graph.

It was observed that the technique was not specifically targeted at HTA. However, as survival analysis is quite common in HTA, it was thought to have enough relevance to be taken to initial development stage.

4.2.8 Graphic 8 – Distribution-based forest plot

4.2.8.1 Background

Forest plots are widely used and accepted in health research to show the outcomes of multiple clinical trials, especially when meta-analysis is conducted. Their use is specifically recommended in the most recent NICE methods guide (National Institute for Clinical Excellence 2004).

The forest plot, developed specifically for use in meta-analysis (Lewis & Clarke 2001), shows the overall direction and weight of evidence in a set of clinical trials (see Figure 4.2 – 14). The squares are positioned horizontally, according the sample mean reported in each trial. These squares are sized according to the weighting given to the trial.* The thin horizontal lines show the degree of uncertainty in the trial, frequently presented as the 95% confidence intervals†.

Forest plots allow for a good overview of the trials in a review. However, this simplification may obscure some of the detail given in a trial report.

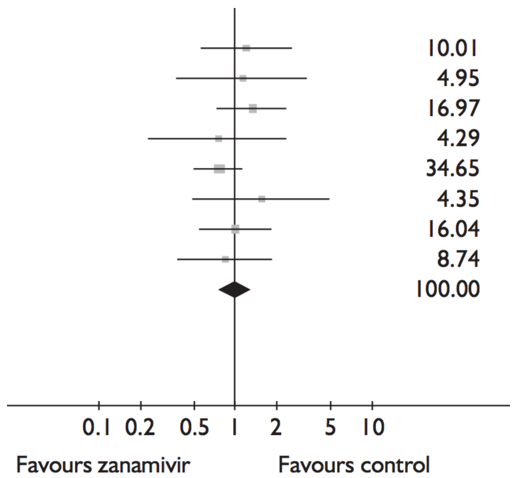


Figure 4.2 – 14
Forest plot (from Burch 2009)

4.2.8.2 Description

This graphic aims to show more detail on how certain and reliable the trials in a meta-analysis were. The forest plot's familiar squares with error bars are replaced by a vertically centred probabilistic distribution. The idea is that the more likely it is that the true value falls at a particular level, the wider the shape at that point. A combined summary of these distributions replaces the customary summary diamond at the bottom of the diagram (see Figure 4.2 – 15). The weighting of each trial, displayed with the size of the squares in a traditional forest plot, is instead displayed using colour or greyscale shading.

No trial can predict the outcome of a medical intervention in an entire population from a sample with 100% accuracy. Using the point-and-

* This weighting can be as simple as how many people were in the trial, but is frequently a more complex synthesis of size, study design, methodological quality and other factors.

† A statistical measure, giving a range of values and a stated degree of certainty. Commonly, this range represents an area within which there is 95% probability that the true population mean may be found, as estimated from a sample of this population (Field, 2005)

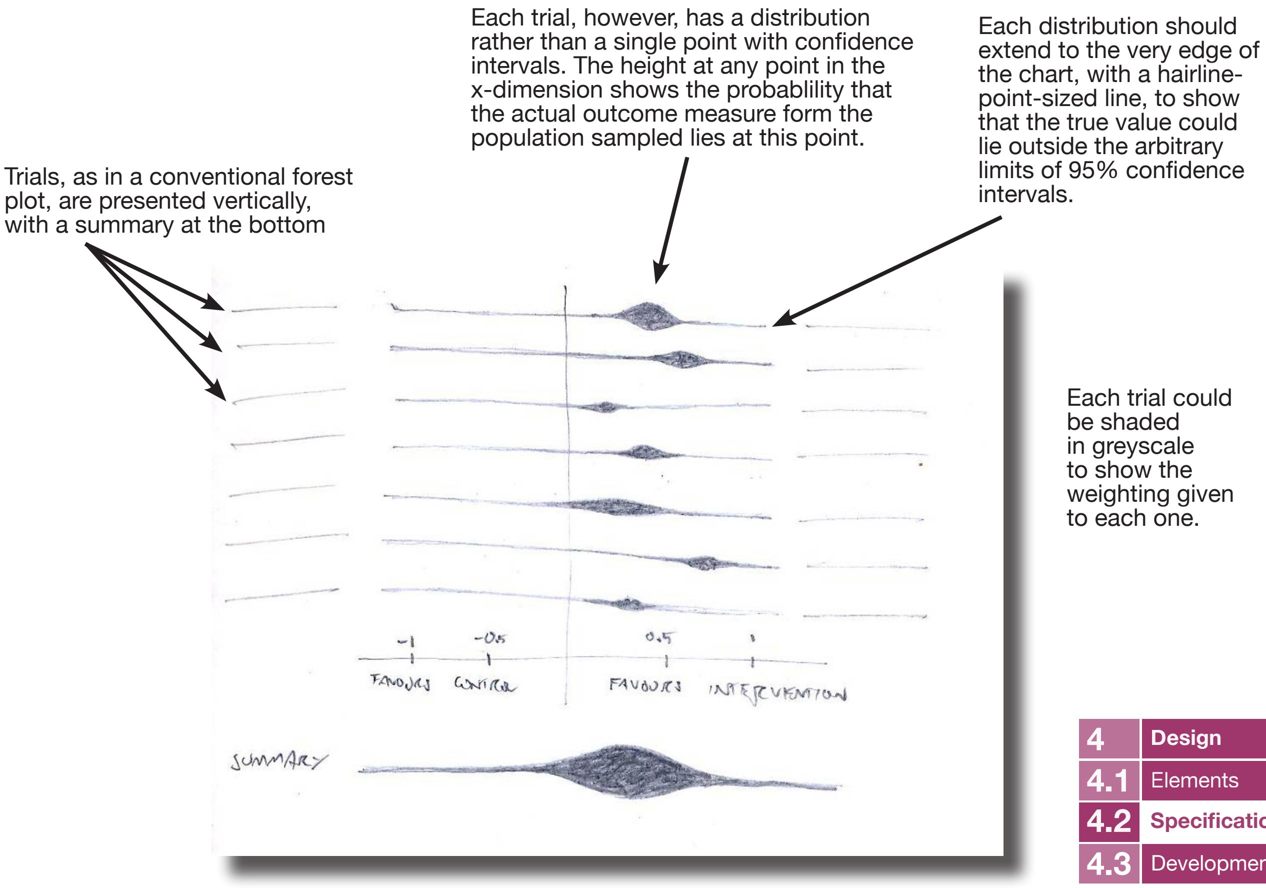


Figure 4.2 – 15
Distribution-based forest plot

whisker form of the forest plot, several different probabilistic distributions could possibly be represented with the same point-and-whisker form. The representation of the result of a trial as a single point with 95% confidence could suggest an even distribution of probability when this is not, in fact, the case. Not all users of TAR reports are likely to be experienced enough to be able to correctly interpret the simplified statistical information contained in a forest plot.

4.2.8.3 Expert critique

Even if it were possible to create such a graphic from the information available, it was thought that people would still want to see key points, like the sample mean and 95% confidence intervals on the graphic.

However, the more important issue was that the display might not be viable in terms of the statistical information available from trial reports. In the frequentist statistics employed in HTA, the distributions used are uniform, so no information is given about where the population mean is likely to appear between the 95% confidence intervals. Bayesian statistics might provide this kind of information, but as this is not commonly employed in HTAs at the time of writing (Ades et al. 2006), this graphic was not chosen to move to the initial development stage.

4.2.9 Graphic 9 – Search strategy link diagram

4.2.9.1 Background

The need addressed by this graphic was identified by an information scientist, who highlighted a specific problem that they had had in their HTA work. They wanted to compare the search strategies used in five different (but related) systematic reviews. Each of the reviews used a different search strategy, with between 20 and 42 search strings used. They found it quite difficult to get any kind of overview of this information quickly.

4.2.9.2 Description

The individual search terms used are represented by small circles, which are placed into one or more of five larger circles, one for each systematic review. Each term is then linked to any other use of that term in another review with a thin line (see Figure 4.2 – 16).

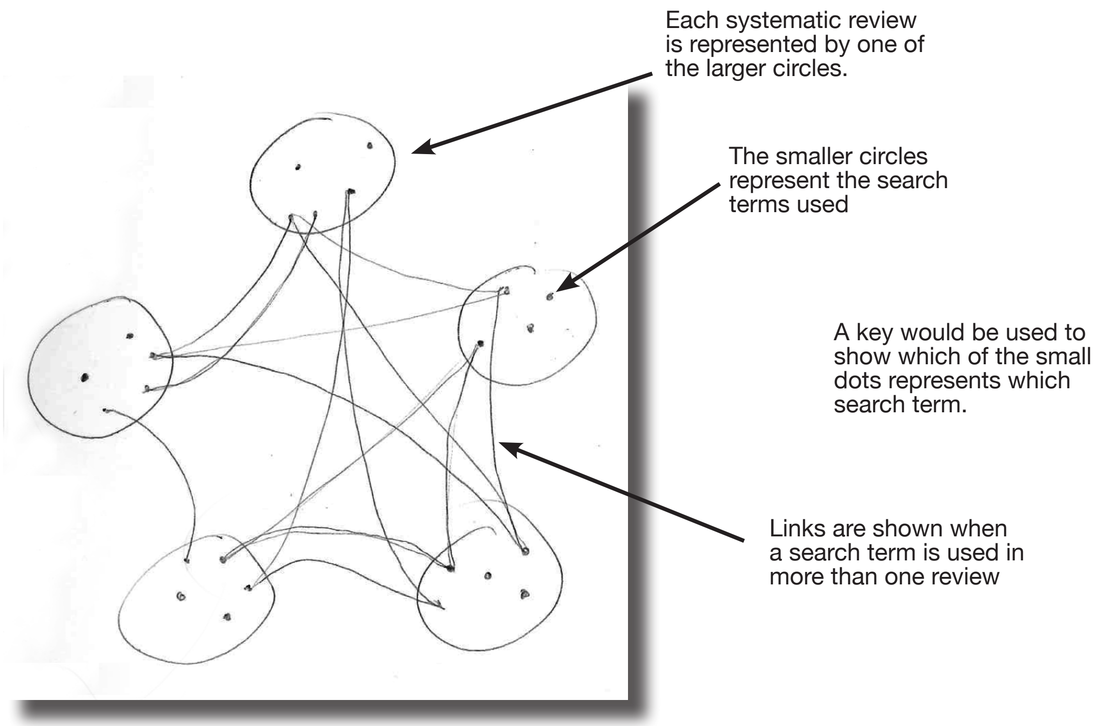


Figure 4.2 – 16
Search strategy link diagram

There are many ways in which this information could be displayed. A more straightforward, and less graphical approach, would be to construct a large table, with 93 rows for the different search terms, and 5 columns for the five systematic reviews. However, apart from the very tall narrow table being difficult to present on screen or printed media, it would be difficult to get an overview of each systematic review individually. The information graphic presentation not only gives a count of the search terms used, but also shows easily those terms not used by any other reviews (those with no links). It could be used to give an overall feel of the search strategies used, to someone that was not familiar with the data, in a short time.

4.2.9.3 Expert critique

While this graphic was thought to be of some interest, it didn't go far enough for the panel's liking. It was noted that the graphic does not give any information about how many papers were found with each search term, and how many duplicates were found, which are often key aspects of a search strategy. If presented with a graphic like this, it might be natural to assume that the fullest circle would be the best search strategy. However, without linking this information to the results returned, it is hard to know if that is the case. A large set of search terms might lead to too many results, which might provide the same papers after exclusion, but less efficiently.

Scalability and generalisation were also discussed. In this case, five systematic reviews and 93 search terms were used. It might struggle to represent 20 reviews, or 1000 search terms, however.

Despite some reservations from the panel, this graphic was chosen to be taken to initial development stage, to see how clear the graphic would be with real data incorporated.

4.2.10 Graphic 10 – Individual patient display for discrete event simulation

4.2.10.1 Background

Discrete event simulation can be a useful way of producing a model of the possible results of adopting a health intervention. While the Markov models more commonly used in HTA simulate an entire cohort of people together, discrete event simulation tracks individual units (in this case patents) through the model. For example, individual patients might move between well and unwell states, and back again. In a discrete event simulation model, each individual patient could be traced through the model, and their history recorded, whereas the Markov model would only show how many people were in a given state at any time. However, this can make discrete event simulation models more complex, and create the need for some kind of overview to aid understanding.

Each individual virtual patient in the model contributes to the overall results, but it can be difficult to see if any individual people are unduly affecting the results. Showing the values produced by different patients in the simulation might also enable a decision-maker to look individual patients' journeys through the model, helping them to decide whether patient pathways are realistic, in terms of their clinical experience (model validation).

4.2.10.2 Description

This technique is likely to produce quite a complex visual effect, and is therefore probably better suited to an interactive (screen) display where selective focussing and/or animation could reduce the level of information displayed at any one time. Key findings could, however, be entered into reports, in a similar way to the graphic described in Chapter 4.2.3 – parallel coordinates for probabilistic sensitivity analysis.

In the display, circles are used to represent individual patients, the costs they have so far incurred, and the associated utility scores (see Figure 4.2 – 17). All of the simulated patients' journeys through the different health states could be shown at once, but it might be most useful to highlight individual cases of interest for discussion, perhaps fading the other patients to a light grey colour.

One great problem facing such a display, especially in the case of 1000 or more simulated people, is that occlusion is likely to affect the lines leading from one state to another. This would be a particular problem for people placed in the representation in a position diagonally aligned with another person, making the same transition simultaneously from one state to another. This might be partially addressed by ensuring that some extra space is left either horizontally or vertically, so that the states are not directly aligned with each other.

By showing the costs, utilities and pathways using a graphical technique such as this, unusual cases can be highlighted and analysed, and the face validity of the model can be assessed by clinical experts to validate the model.

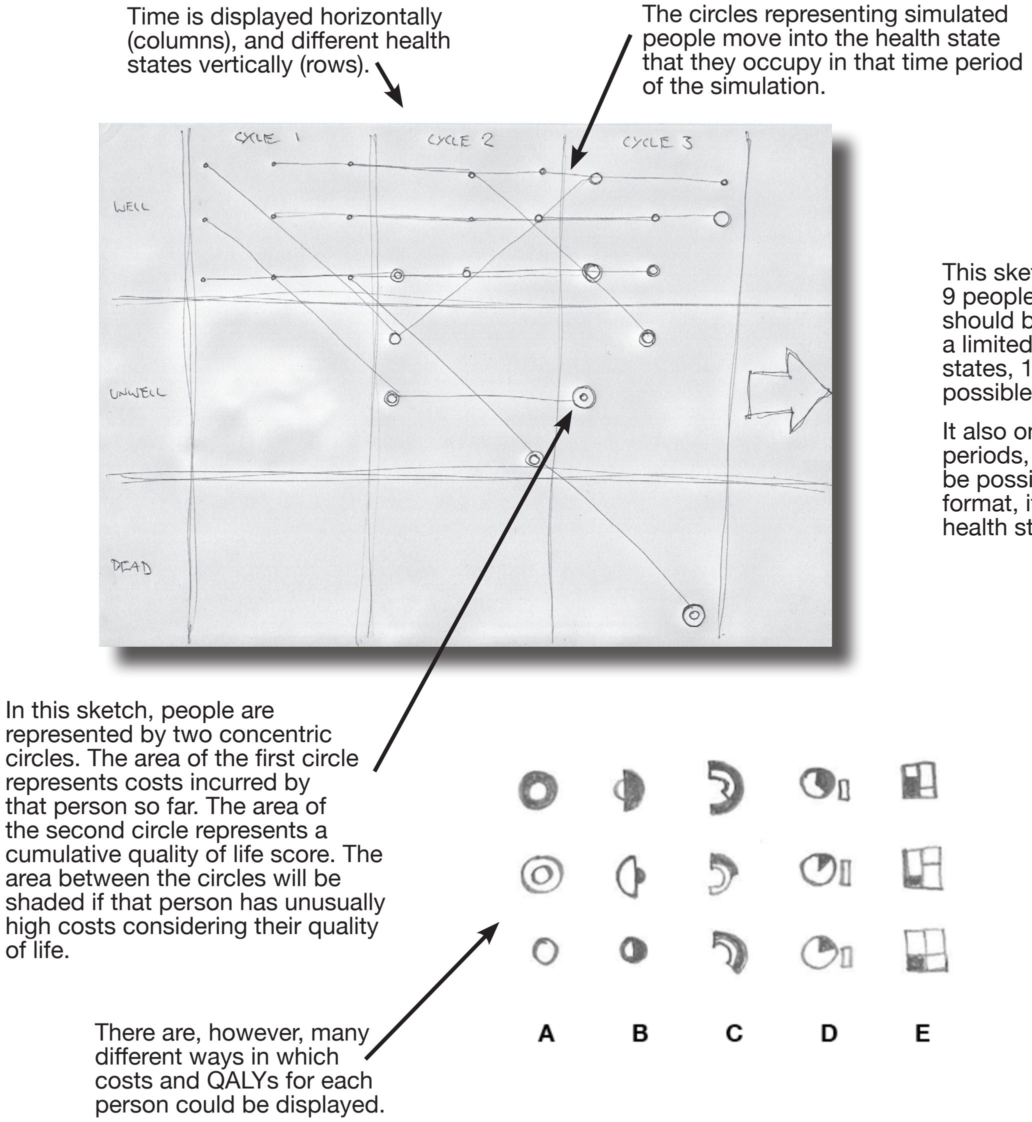


Figure 4.2 – 17
Individual patient display for discrete event simulation

4.2.10.3 Expert critique

The volume of data that would need to be presented could be a significant problem for this kind of presentation. It was noted in the discussion that discrete event simulations often use thousands of people, and track them over hundreds of cycles. The resolution needed to display that many people is probably unrealistic, even in a simple three-state model.

The graphic would also enjoy limited applicability, as discrete event simulation is rarely used in HTA. With other, simpler, graphics specified for more commonly used techniques, this graphic was not thought to be worth the effort it would require.

On top of these substantial criticisms, it was mentioned that simpler animated techniques are already frequently employed in displaying discrete event simulations, and widely available in discrete event simulation software. The graphic specified here is a slightly different way of presenting these results, so it does have marginal interest, but certainly not enough to justify the technical challenge of its development. It was decided that the development of this graphic should not continue in the next stage of the project.

4.2.11 Potential colour variations

HTA reports are currently distributed to NICE decision-makers in printed form, using only black ink on white paper. They are also published in the HTA journal, which uses black ink only. However, electronic versions of the reports are hosted on the NICE website, in PDF format. This file format is ideal for the presentation of static information graphics, as it allows vector-based* coloured representations. If colour printing became possible in future, or the online PDF became the dominant format for decision-makers to view the reports, coloured variations of these specified information graphics would be possible.

* Vector-based files, such as PDFs, can be enlarged and resized, as they are recorded as a series of mathematical functions, rather than on a pixel-by-pixel basis, as in bitmap (or raster) file formats such as JPEG and TIFF).

In graphic 2 (bubble chart for two-way sensitivity analysis), for example, a different colour could be used for positive and negative changes to the ICER. Instead of a medium-sized circle representing no change, it could be shown as a circle of area 0 (and therefore invisible). This would mean that picking out a large negative change in the ICER would be much easier, as it would appear as a large circle in a different colour to the positive changes, rather than having to look for very small circles. This use of colour for categorical distinction could also be useful in graphic 5 (sankey diagram for overview of Markov model), to distinguish between the different model states, and link these to the bar charts showing costs and QALYS accumulated above and below the main part of the diagram.

Colour can also be used to present another dimension of data, using a colour scale. This might require high quality, carefully colour-calibrated printing, or high quality digital displays. However, the extra dimension of data could be very valuable. In the case of graphic 7 (survival synthesis bubble chart), the bubbles could be coloured with a scale, which would represent the hazard ratio† at that time point in the trials.

† A statistical measure of the hazard, or risk of an event, with reference to an explanatory variable – in this case the risk of death with reference to which of two interventions was received.

4.2.12 Potential animated variations

The use of time-based media would allow for animated information graphics, which could be distributed as video files. These could be given to decision-makers as a streaming link over the internet, or provided with the reports on DVD.

This would allow graphic 10 (individual patient display for discrete event simulation) to show what happened to people over time. It might also be used for graphic 5 (sankey diagram for overview of a Markov model), allowing each cycle of a very lengthy model to be shown scrolling past, and showing the gradual build-up of costs and QALYS.

4.2.13 Potential interactive variations

Digital media also offer scope for interactivity, whether this is used to allow the user to limit the amount of data shown, or to enable them to enter their own values. Interactive displays might be designed with programs such as Flash or Processing, for example. The resulting Flash player files could be distributed using the internet, and Processing sketches can be converted into Javascript. Alternatively, such interactive graphics could be shown at committee meetings using data projectors. This added interactivity could enhance several of the graphics.

Graphic 1 (small multiple techniques including Sankey diagrams for overview of clinical effectiveness data) could be designed so that the user could sort the display by clicking on one or more of the headers, in a similar way to a spreadsheet. This would allow them to see, for example, the trials with the lowest ages or the highest number of people, and could reveal patterns in the data.

Another possibility for interactive graphics would be to provide more detail when highlighting a specific element, perhaps by moving a mouse pointer over it, clicking, pressing on it, or pointing in some other way. In graphic 7 (bubble chart for survival synthesis), a user could highlight a circle, and see more information about the characteristics of the trial from which that data point comes – trial size, design and methodological quality, for example. The display could also show the viewer the more familiar Kaplan-Meier curve* for this trial, and show where the data points that they have highlighted would be on

* See Chapter 4.2.7.1

this graph. In graphic 8 (search strategy link diagram), the user could highlight specific search terms from a list, and see them picked out in the link diagram. Also, they might activate an entire systematic review circle, and see only the links to other reviews highlighted, and only the search terms used highlighted.

4.3 Development

After discussion, five of these specified information graphics were selected for further development. These were produced in a form that could be presented in a technology assessment report, using real data from past reports to test the concept of the information graphic. This section (4.3 – development) details the results of this practical investigation of the merits of these five information-needs-based designs. As such, it can be regarded as a modest application of the 'practice-based research' discussed in Gray and Malins (Gray & Malins 2004).

This 'development' process forms the third phase of the suggested design process for information graphics in HTA (see Figure 4.3 – 1).

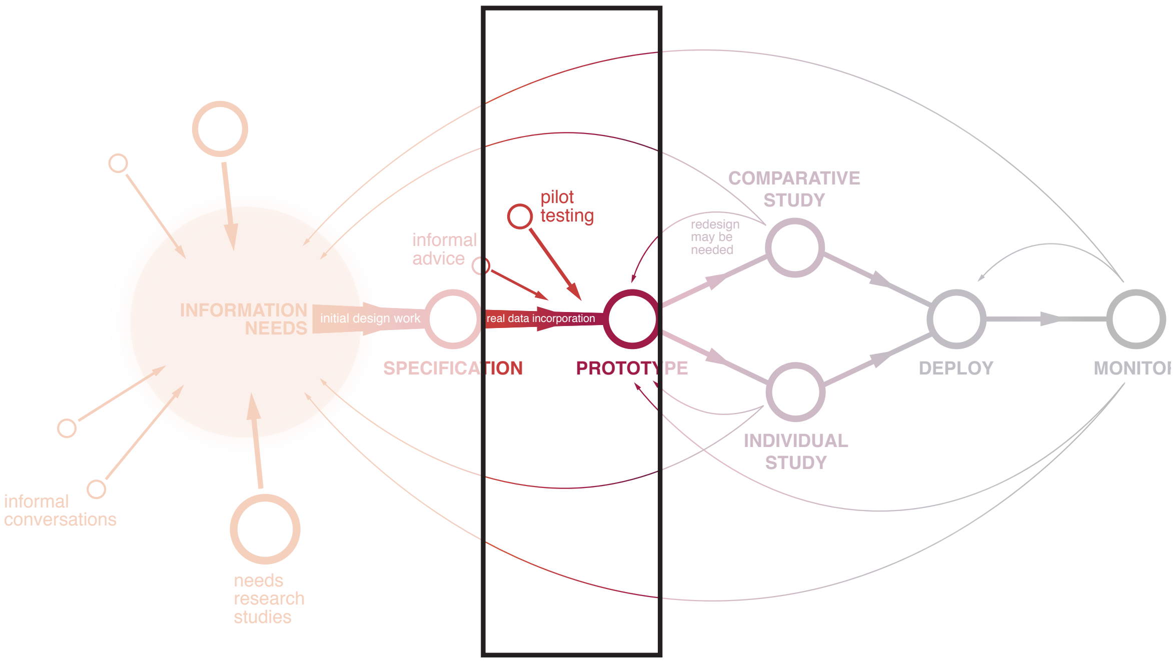


Figure 4.3 – 1
The phase of the proposed design process detailed in Chapter 4.3

4.3.1 Graphic 1 – Small multiple techniques including Sankey diagrams for overview of studies in a systematic review

(See Chapter 4.2.1 for specification)

An initial version of this graphic was created with eight of the trials included in the original report. This is shown in Figure 4.3 – 2 and Figure 4.3 – 3 over the following two pages. As this would be a multiple page document, Adobe InDesign software was used for the layout. The small multiple Sankey diagrams for the study design and size were produced in Adobe Illustrator, saved as EPS files, and imported into the master document in InDesign. The values for widths, lengths, and positions of the graphics were calculated in Apple's Numbers spreadsheet program, although any spreadsheet could have been used. The values were copied and pasted into the appropriate boxes in InDesign and Illustrator to set the position and size of the various data elements on the page.

There were several changes to the graphic in this version, which came from the comments received on the specification, and in response to design challenges raised by the application of real data:

- The UK trials were picked out with a representation of the shape of the island, as these are likely to be more relevant to the context of NICE decisions.
- Some of the trials did not report mean age. This was simply addressed by removing the vertical line in these cases, leaving an impression of the range of ages used, without showing a specific mean value.
- The statistically significant outcomes were displayed more appropriately, in the grid showing the outcome measures used. This allowed a single trial to display significance in some outcomes, but not all. Solid black circles were used to display significance, so that these outcomes would stand out against the others.
- In the extra space created by moving the statistical significance, a quality checklist was included, showing some important information about how the trials were conducted.

The new 'quality grid' could be used in at least two different ways. It can give a general sense of the quality of the trials, by judging how black the grid is overall. Alternatively, a viewer can pick out an individual box of interest from

Information graphic techniques for overview

of a systematic review of clinical effectiveness for cochlear implants.

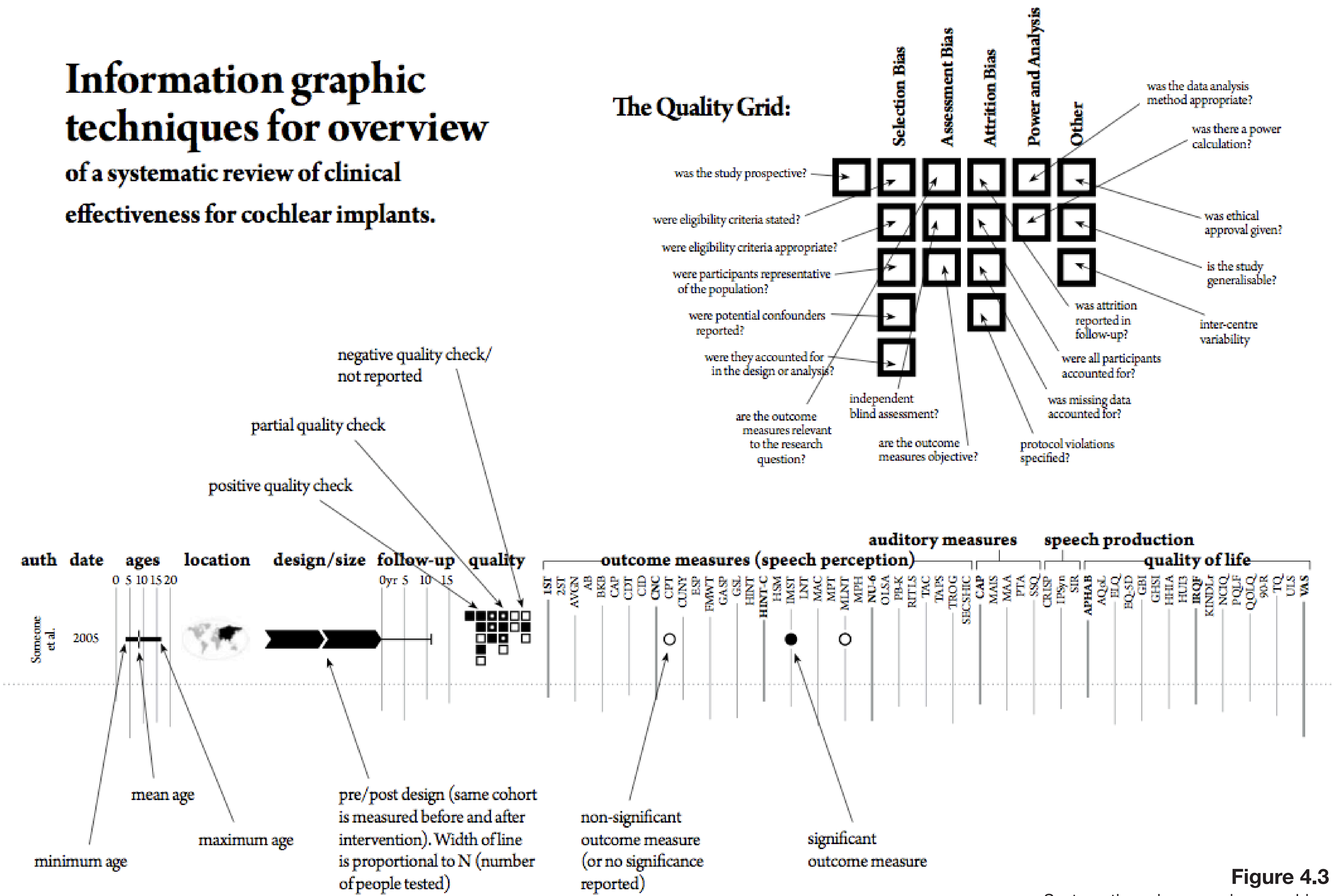


Figure 4.3 – 2
Systematic review overview graphic – key

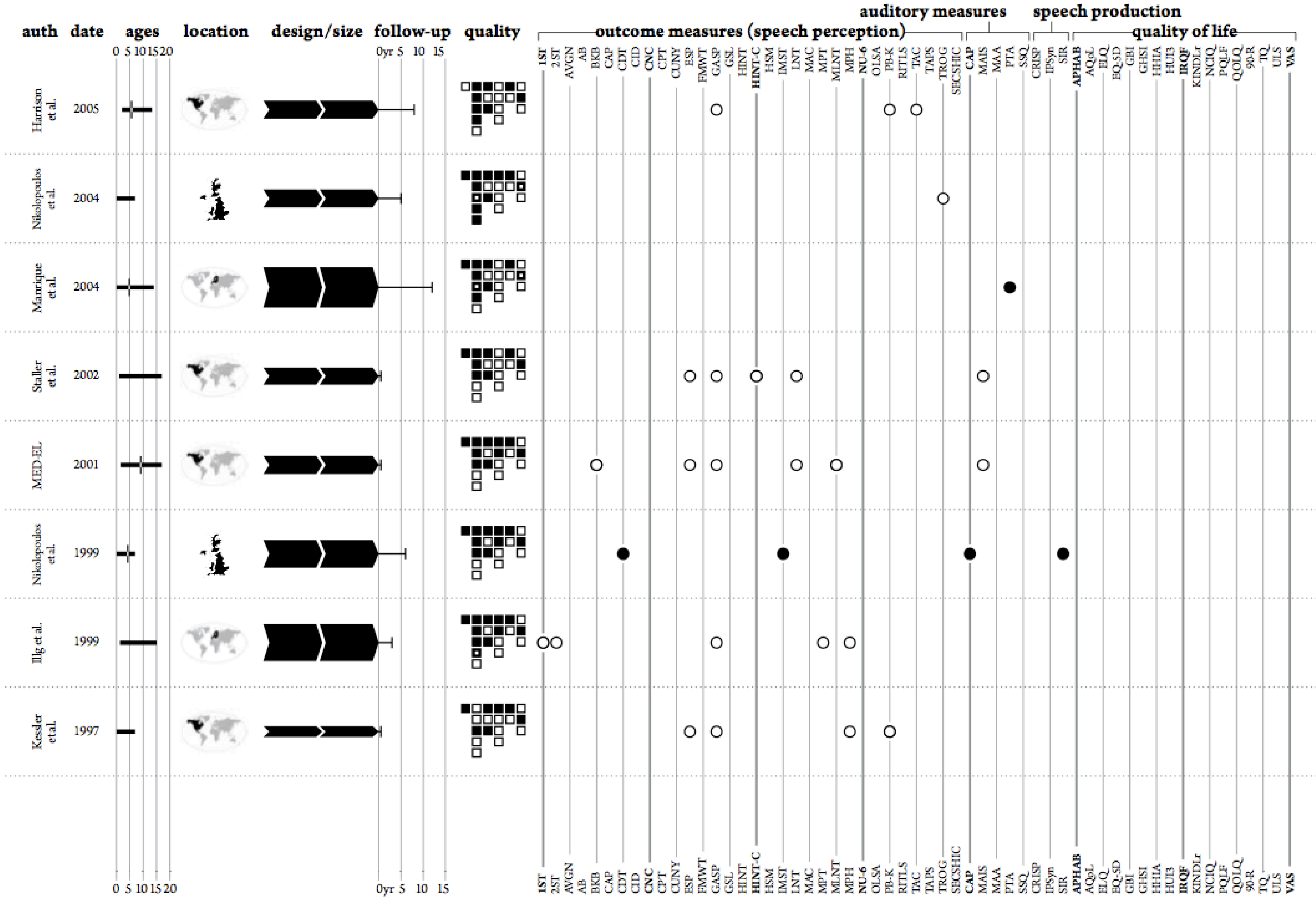


Figure 4.3 – 3
Systematic review overview graphic – page 1

the key, and scan down the document to pick this out. This task should be made easier by the distinctive shape of the grid. The danger with the first of these ways to use the grid is that each box has the same visual weight. However, each box might be more or less important, depending on the study domain, as well as the kind of trial being performed. This was picked out during consultation with the original report author as a weakness with the presentation method, who suggested that some kind of labelling of this grid should be attempted.

They also mentioned that it would be important to have some way of judging the absolute size of the trials. While it was easy to see which trials were larger than others in this set, it was not possible to see whether the trials were large or small overall.

In this case, no trial produced a significant effect that was in favour of the comparator intervention. The original author did ask how that would be identified if it were necessary to do so for another set of trials. This could be shown with a diagonal stripe through the circles, suggesting a negative outcome.

This graphic was relatively easy to develop, taking around two weeks' full-time work. It could be used in almost any HTA report, and seemed to be widely accepted as useful in consultation with the report author and PhD supervision team. It was therefore taken to a further stage of more formal evaluation, detailed in Chapter 5 – prototype test 1.

4.3.2 Graphic 3 – Parallel coordinates for probabilistic sensitivity analysis

(See Chapter 4.2.3 for specification)

The data for the initial development of this graphic came from another PenTAG report, this time on two treatments for Glioma (Garside et al. 2007). This report was chosen, as it was a fairly simple model design, and access was easy to the original Excel file with the values calculated for the PSA. The data from just the analysis of TMZ (temozolomide) vs placebo was used to demonstrate the principle of the graphic. The probabilistic sensitivity analysis of the model in this report used 1000 samples of different variables.

The graphic was created using a free data visualisation program called Orange (from http://www.ailab.si/orange/) to create the parallel coordinates display. This was then exported to Adobe Illustrator for labelling, as an SVG file (a vector format)

The large number of variables in the PSA led to a very complex looking graphic, even with benefit of colour. Figure 4.3 – 4 shows a snapshot of the parallel coordinates, with all simulations which produced an ICER below £35k per QALY highlighted in red. A threshold of £30k per QALY is more commonly used as a metric of interest (Devlin & Parkin 2004), but in this case, only two of the simulations produced an ICER this low.

There was no very clear clustering of the red lines on any single variable. The lower costs seemed to be produced by many different combinations of variables, rather than showing a clear pattern of low costs and high utility scores, as might be expected.

One interesting observation is that the 'simulation number' bar on the left seems visually to cluster at the top and bottom, even though it should be a perfectly linear distribution - it simply shows an ID number for the simulations, running from 1 to 1000. Other bars, such as the 'costs in progressive state, week 1, all arms' bar, may also be more uniform than they appear.

The original modeller suggested that net benefit might be shown as a slider as well as the ICER, which might have different results. This would be technically very easy to accomplish.

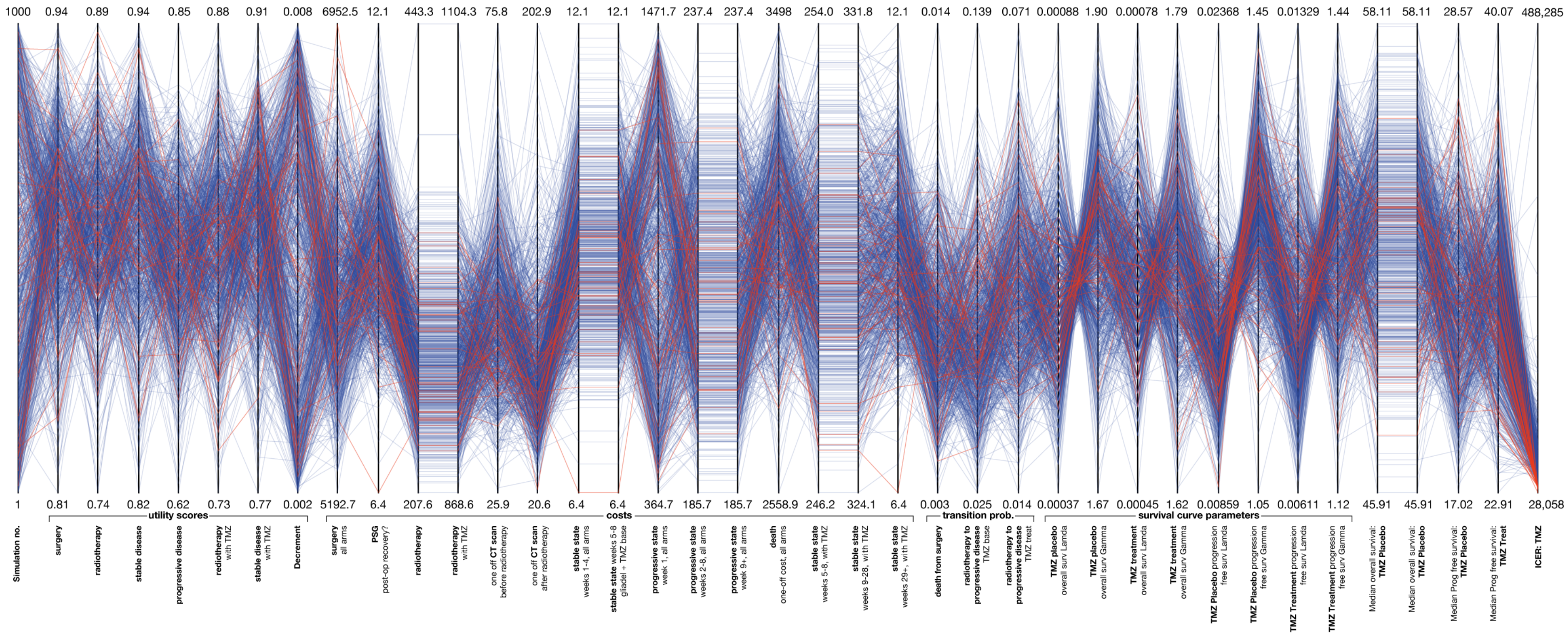


Figure 4.3 – 4
Parallel coordinates for probabilistic sensitivity analysis.

They also suggested that, while this presentation of the graphic might not reveal anything very clearly, it would be possible to reduce the number of variables shown, simplifying the graphic by only presenting key variables that were either very uncertain or had a great effect on deterministic sensitivity analysis. For the deterministic sensitivity analysis, the original report used:

- median survival advantage with TMZ
- utility in stable state with TMZ
- utility in progressive state with TMZ
- cost of TMZ

In addition, there are some variables that seem to stand out from Figure 4.3 – 3:

- survival curve values for TMZ placebo progression free survival
- costs of 'PSG' state (which represents post-operation recovery in the model),
- costs of the CT scan performed after radiotherapy.

It would be possible to create a version of the graphic which only showed these seven variables, which might be easier to interpret.

Unfortunately, it can be difficult to see any dual-influence patterns in parallel co-ordintes displays like this one unless two values that depend on each other happened to be next to each other. While it would in theory be possible to present each pair of variables separately, with 32 variables as we have here, that would require 528 pairwise comparisons. It might be possible to use a display like that shown in Figure 4.3 – 5 to compare the seven variables already mentioned in a pair-wise manner to look for combinations of variables that resulted in low ICERs.

After this initial 'real data' development and feedback stage, too many questions were left unanswered. Scalability issues were becoming apparent even with this relatively simple model. This graphic was abandoned at this point, and not taken to final testing stage.

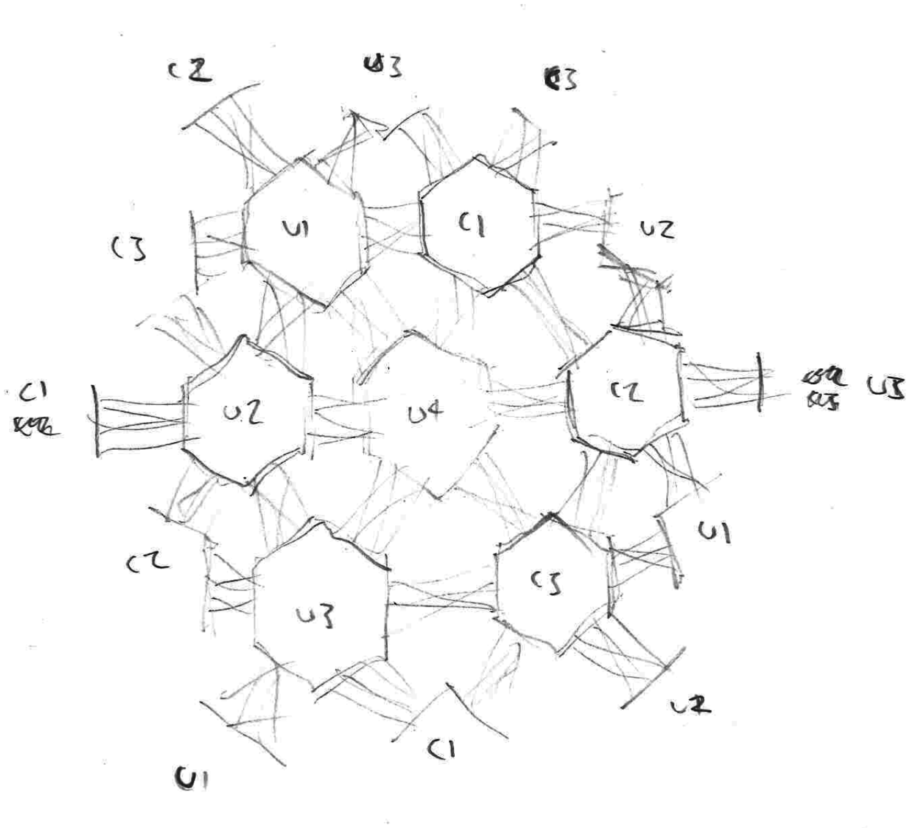


Figure 4.3 – 5
Hexagonal display for comparing seven variables from a probabilistic sensitivity analysis.

4.3.3 Graphic 5 – Sankey Markov overview

(see 4.2.5 for specification)

Data from the glioma model was used again for this graphic (Garside et al. 2007). Figure 4.3 – 6 shows the TMZ arm of the model, for the base case scenario, with the best available data from the systematic review used. The eventual aim with this graphic would be to design specialist software that would output all arms of a model in this format, and also allow deterministic sensitivity analysis to be performed, to show the different sources of the changes in the ICER produced by changing inputs to the model. This prototype version, however, was created using Adobe Illustrator, again copying and pasting values for the position and size of objects directly from a spreadsheet program, in this case OpenOffice (from www.openoffice.org).

The first design challenge was to find a way of displaying the 235 cycles of the model onto an A4 page. Rather than attempting the circular format suggested in Figure 4.2 – 7, the model was broken into 10 26-week blocks. As the first few

Temozolomide treatment arm

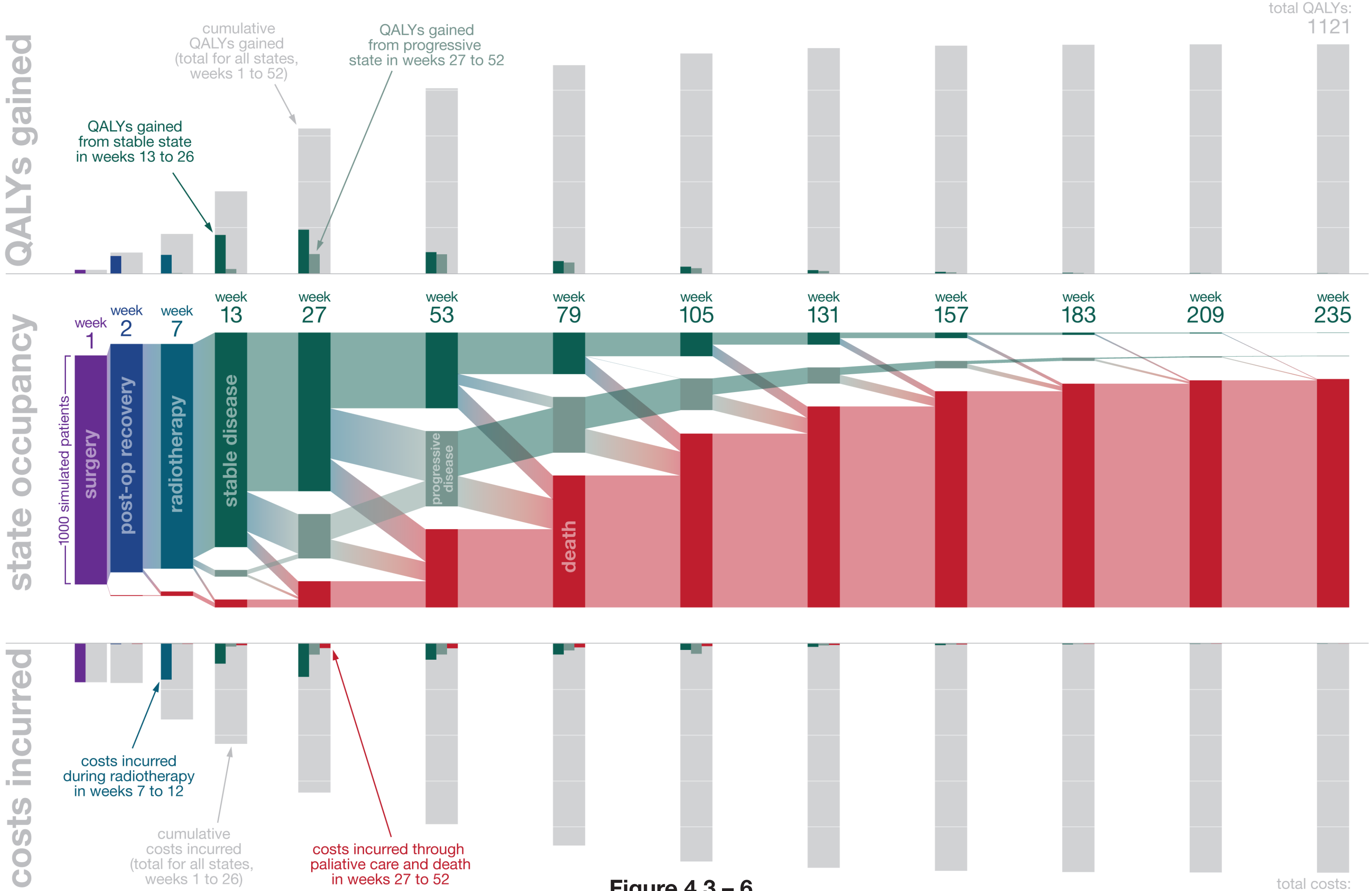


Figure 4.3 – 6
Sankey Markov overview, model of temozolomide for the treatment of newly-diagnosed high-grade glioma

weeks were quite complex, and involved several different procedures, the first of these blocks was further broken into four different parts. Rather than using average values over the 26 week blocks, 'snapshots' of the number of people in each state in individual cycles were used. This meant that the number of transitions between the states in the cycles between one snapshot and the next could. In fact, this did not produce a 100% accurate result, as the number of transitions did not quite sum to 1000. However, the actual values were very close, and converged on the expected values towards the end of the model. The reason for this is currently unknown, but may be a rounding error. However, it does not appear to be a great problem for this model, which is quite simple. The method does seem give a reasonable idea of the number of transitions in this case, but caution should be advised in a more complex model.

This graphic seemed to give a good overview of the source of the costs and QALYs in this model. The colouring of the different states helps to relate the costs and QALYs produced by the model to the health states in which they are produced. It would, however be quite difficult to show this in a black and white, printed report. It might be accomplished with good labelling, but would perhaps not be as clear.

One of the flaws in the graphical presentation is that it does not show the incremental gains or losses between different arms in the model. Instead, a different graphic would needed for every arm. While it would be easy to show the incremental gains and losses in the bar charts above and below the main section of the graphics, the state occupancy section presents a greater challenge. It might, in the end, be of more use for deterministic sensitivity analysis. It would be relatively easy to highlight the differences made from the base case in various different scenarios, but again, changes to the state occupancy would be difficult to show. Bars with about a quarter of the width of the current cycle bars might be included just inside the left of each cycle shown in the graphic, to show what the base case was like. This would require careful explanation, however, and some degree of learning on the part of the viewer would be necessary.

There seemed to be many possibilities for further development of this graphic. Also, it dealt with a different part of the HTA process to graphic 1. Therefore, this graphic was also chosen to go to the final user testing stage, which is detailed in Chapter 6 – Prototype test 2.

4.3.4 Graphic 7 – Survival synthesis bubble chart

(See Chapter 3.4.2.7 for specification)

This graphic was developed into an interactive version, which could be presented online, in which data about each trial appeared when the user put their mouse pointer over the graphic*. A static screenshot is provided in Figure 4.3 – 7. The circles representing the percentage of surviving participants could be coloured according to the hazard ratio calculated for that time point. The interface was designed using Adobe Flash, using the built-in programming language Actionscript 3 to define the size and position of the circles.

* This is hosted at www.pcmd.ac.uk/infographics/GRsurvival at the time of writing.

The response from the researcher who suggested the area of application for the graphic was that it was generally useful. However, the hazard ratio information didn't come through strongly enough, and it would be difficult to get the graphic published in a journal in its current form.

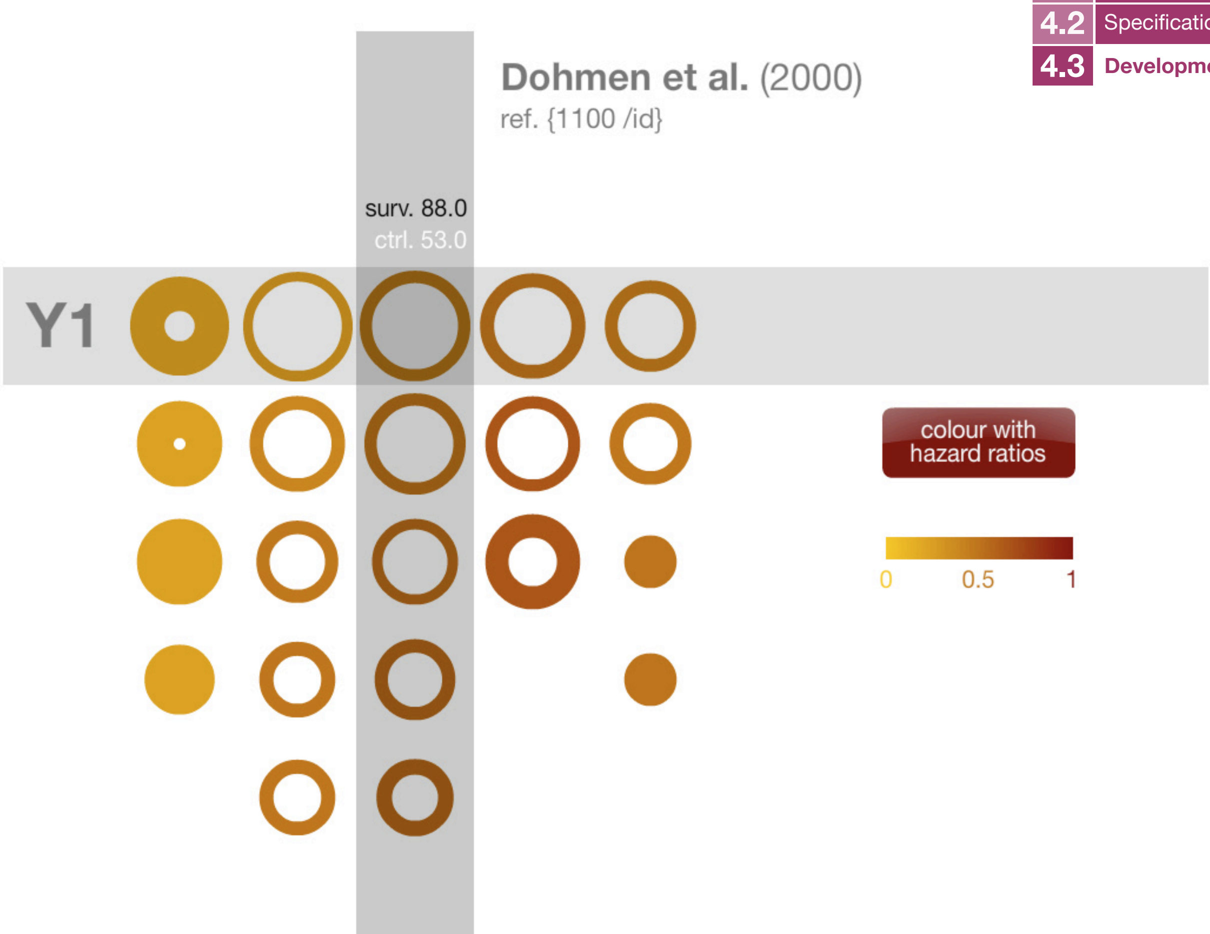


Figure 4.3 – 7
Initial version of the interactive display.

A new version of the Flash interface was designed,* shown in Figure 4.3 – 8 as a static screenshot. This version attempted to bring out the hazard ratio information by changing the colour scale from the original 0 to 1 to a 0.2 to 0.9 scale, that would include all values in the data set presented in this instance. The scale was also moved and enlarged, for clarity. Also, a hazard ratio 'marker' was placed on the colour scale, which would indicate the exact hazard ratio value of the point that had the mouse over it.

* This is hosted at www.pcmd.ac.uk/infographics/GRsurvival2 at the time of writing.

The graphic was at this point presented at the monthly seminar of the PenTAG research group at the University of Exeter.

One of the feelings expressed by the researchers present was that the circle display, or bubble chart, was too different from the more usual 'Kaplan-Meier' curves (Kaplan & Meier 1958) to be understood in its present form. One possible strength of the interactive technique, however, is that it could be used to also show the survival curves in their more familiar Kaplan-Meier form, indicating where the data points used in the selected circle appear on the curve.

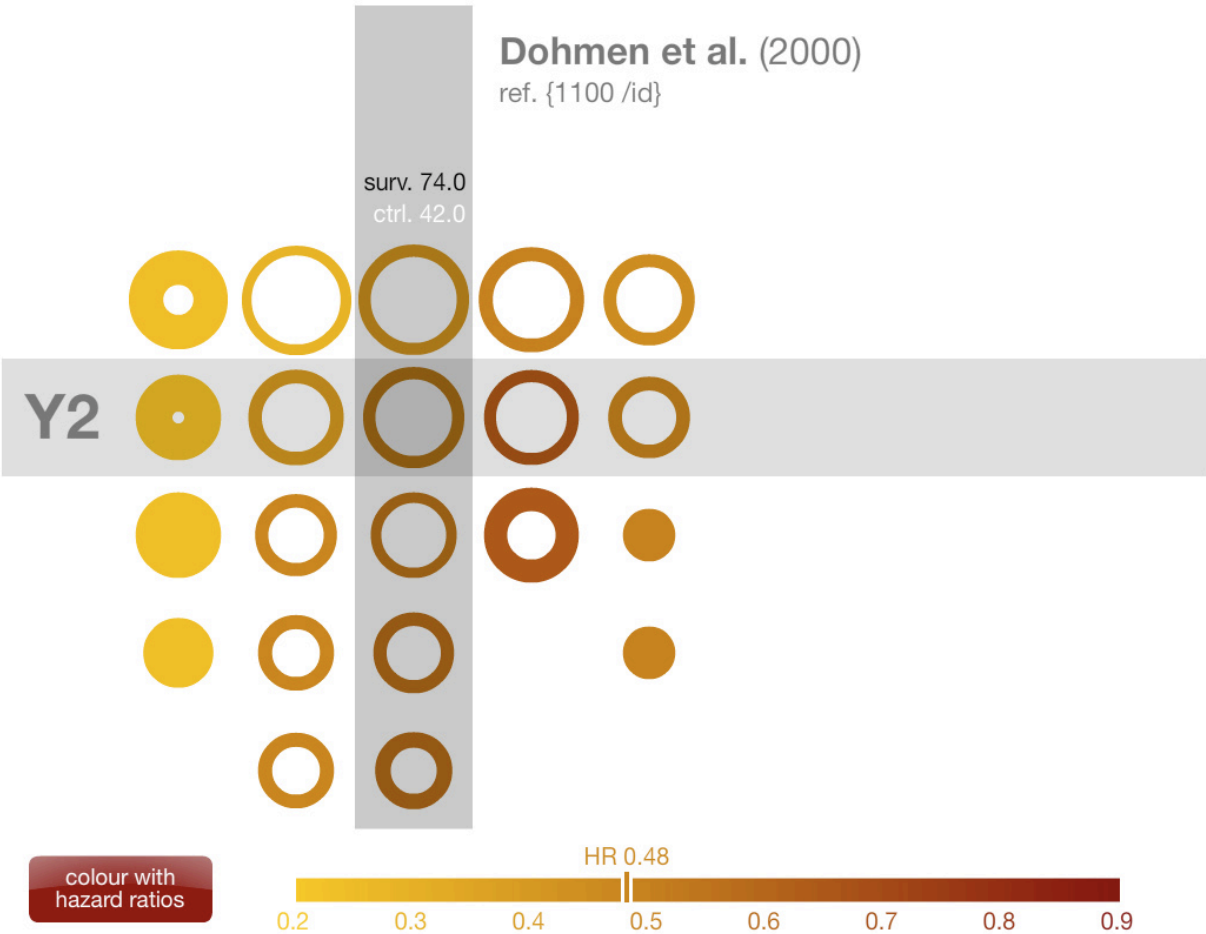


Figure 4.3 – 8
Second version of the interactive display.

More information on the trials was also requested when the mouse was pointed to the circles, such as the trials' size and location.

While this was certainly thought to be a promising method of displaying survival information, it was not chosen as a candidate for final testing. Its reliance on an interactive display puts it at a disadvantage compared to the immediately useable graphic 1, which also focussed on presenting clinical effectiveness data.

4.3.5 Graphic 9 – Search strategy link diagram

(See Chapter 3.4.2.9 for specification)

This graphic was also developed further, using Adobe Illustrator to present the search strategies of the five systematic reviews.

Initially, the search terms used were placed in an unordered way within the circles representing the five reviews. Curved lines were used to link the search terms that were used by more than one review. However, this quickly became confusing (Figure 4.3 – 9).

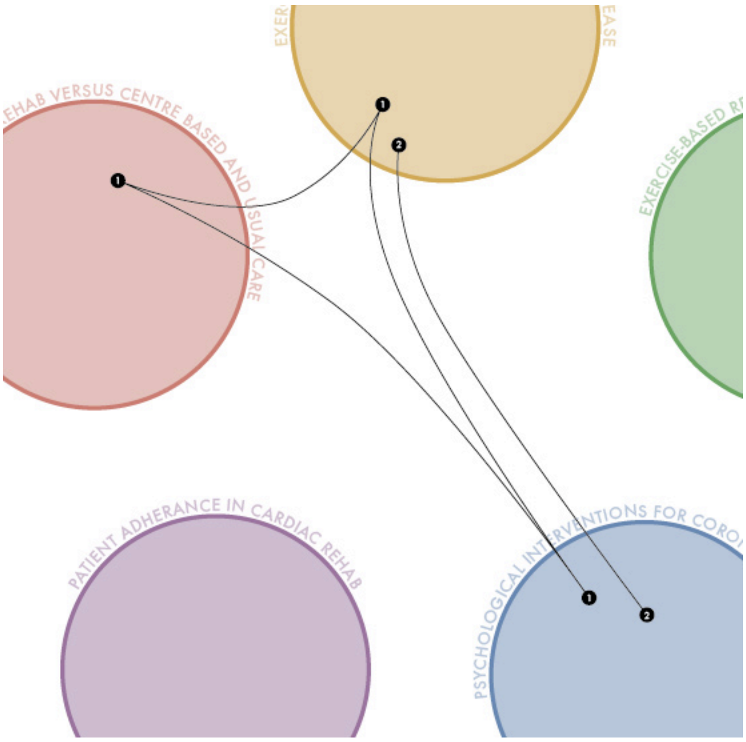


Figure 4.3 – 9
Curved link lines, unordered small circles.

A version with straight lines seemed better, but the unordered placement would clearly be able to obscure the message of which reviews had strong similarities, when the lines were so prone to overlap (Figure 4.3 – 10).

A more ordered layout, with the 93 search terms used arranged in a 10 x 10 grid, was more useful. The large circles representing the systematic reviews had to be adjusted positionally to avoid any two of them sharing the same x or y position, which would cause occlusion (Figure 4.3 – 11).

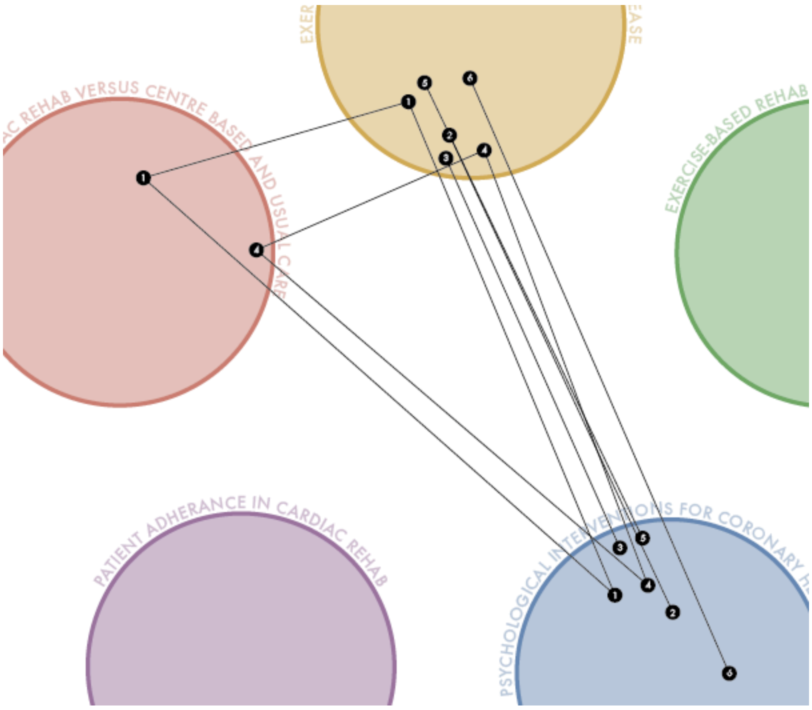


Figure 4.3 – 10
Straight link lines, unordered small circles.

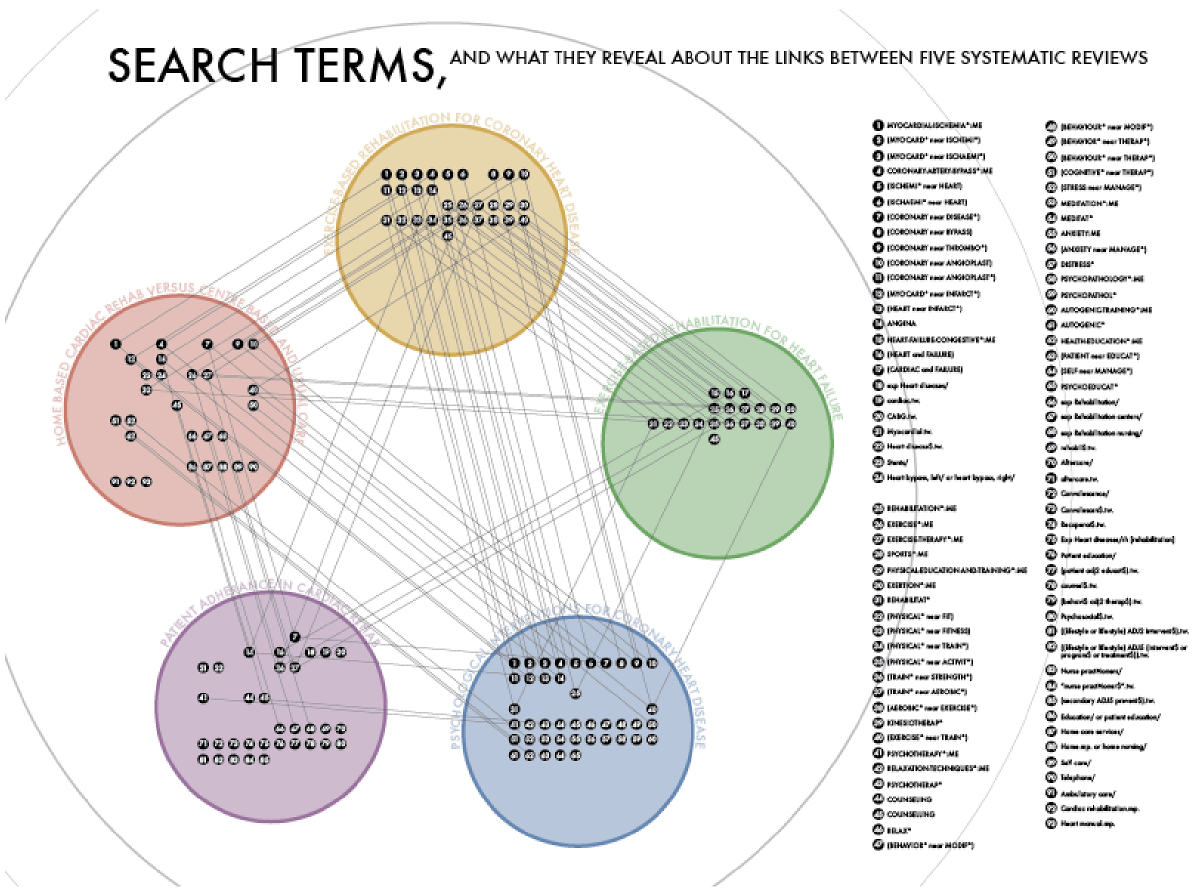
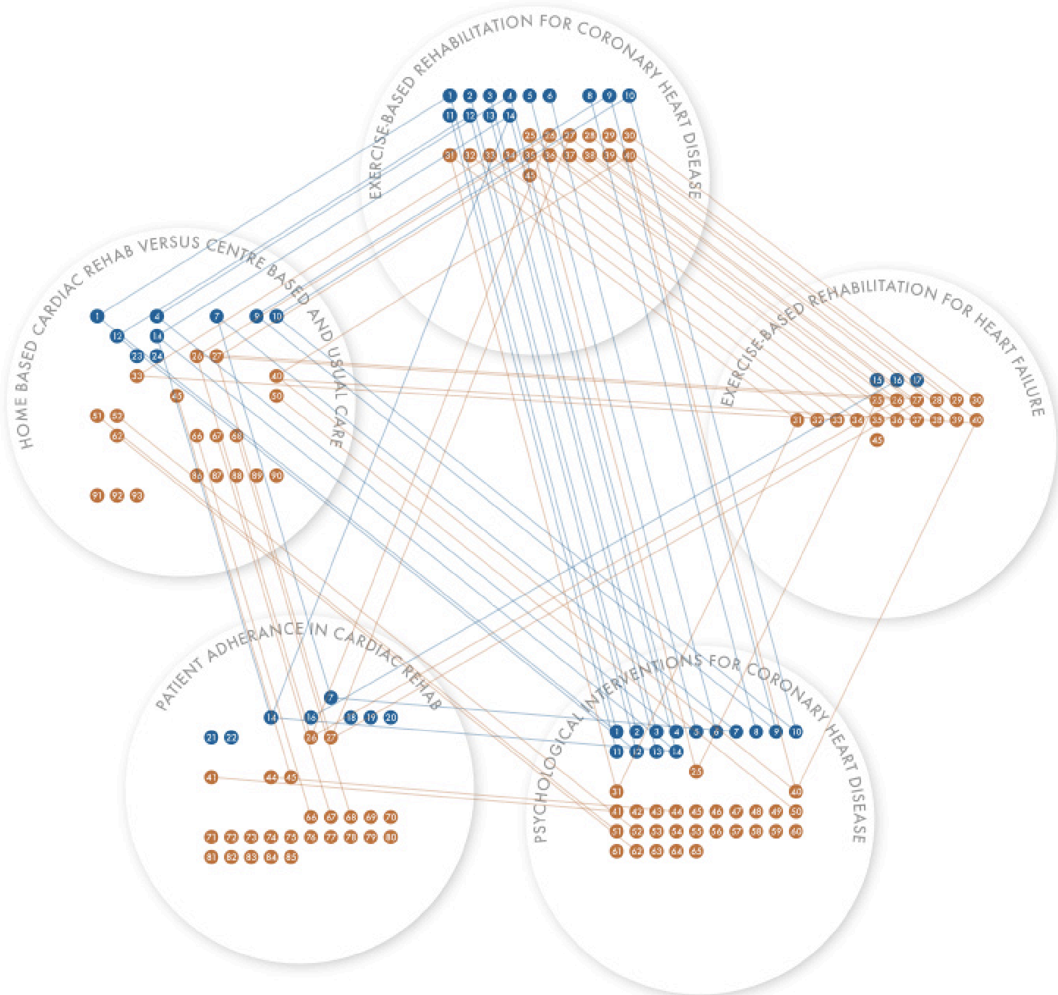


Figure 4.3 – 11
Straight link lines, ordered small circles.

At this stage, the information scientist who provided the data mentioned that it would be helpful to know whether the search terms dealt with the population or the intervention of interest. As it happened, the colour coding of the circles representing the systematic reviews had become redundant, as they were already identified by their position. Therefore, a final version was produced, which used different colours to categorise the search terms used in the reviews instead (see Figure 4.3 – 12).

This revealed strong relationships between two of the pairs of reviews. 'Exercise-based Rehabilitation for Coronary Heart Disease' shared many 'intervention' search terms with 'Exercise-based Rehabilitation for Heart Failure', as might be expected. 'Exercise-based Rehabilitation for Coronary Heart Disease' also shared many 'population' search terms with 'Psychological Interventions for Coronary Heart Disease', again unsurprisingly. However, 'Patient Adherence in Cardiac Rehab' seemed to share surprisingly few search terms with 'Home Based Cardiac Rehab Versus Centre Based and Usual Care', particularly in terms of the population searched for.

While this graphic was starting to produce a few interesting results, there were potential scalability issues, as it might be difficult to use this graphic in reviews with many more search terms. Also, the expert critique revealed that it would only be truly valuable if it could include data on how many papers were returned by the searches, or how the search terms were linked with 'and' / 'or' terms. It was becoming clear that the graphic would need a complete redesign to be able to display this kind of information. Therefore, the graphic was not taken to the final testing stage.



SEARCH TERMS,

AND WHAT THEY REVEAL ABOUT THE LINKS BETWEEN FIVE SYSTEMATIC REVIEWS

POPULATION

1 MYOCARDIAL-ISCHEMIA*:ME
2 [MYOCARD* near ISCHEMI*]
3 [MYOCARD* near ISCHAEMI*]
4 CORONARY-ARTERY-BYPASS*:ME
5 [ISCHEMI* near HEART]
6 [ISCHAEMI* near HEART]
7 [CORONARY near DISEASE*]
8 [CORONARY near BYPASS]
9 [CORONARY near THROMBO*]
10 [CORONARY near ANGIOPLAST]
11 [CORONARY near ANGIOPLAST*]
12 [MYOCARD* near INFARCT*]
13 [HEART near INFARCT*]
14 ANGINA
15 HEART-FAILURE-CONGESTIVE*:ME
16 [HEART and FAILURE]
17 [CARDIAC and FAILURE]
18 exp Heart diseases/
19 cardiac.tw.
20 CABG.tw.
21 Myocardial.tw.
22 Heart disease$.tw.
23 Stents/
24 Heart bypass, left/ or heart bypass, right/

INTERVENTION

25 REHABILITATION*:ME
26 EXERCISE*:ME
27 EXERCISE-THERAPY*:ME
28 SPORTS*:ME
29 PHYSICAL-EDUCATION-AND-TRAINING*:ME
30 EXERTION*:ME
31 REHABILITAT*
32 [PHYSICAL* near FIT]
33 [PHYSICAL* near FITNESS]
34 [PHYSICAL* near TRAIN*]
35 [PHYSICAL* near ACTIVIT*]
36 [TRAIN* near STRENGTH*]
37 [TRAIN* near AEROBIC*]
38 [AEROBIC* near EXERCISE*]
39 KINESIOTHERAP*
40 [EXERCISE* near TRAIN*]
41 PSYCHOTHERAPY*:ME
42 RELAXATION-TECHNIQUES*:ME
43 PSYCHOTHERAP*
44 COUNSELING
45 COUNSELLING
46 RELAX*
47 [BEHAVIOR* near MODIF*]
48 [BEHAVIOUR* near MODIF*]
49 [BEHAVIOR* near THERAP*]
50 [BEHAVIOUR* near THERAP*]
51 [COGNITIVE* near THERAP*]
52 [STRESS near MANAGE*]
53 MEDITATION*:ME
54 MEDITAT*
55 ANXIETY:ME
56 [ANXIETY near MANAGE*]
57 DISTRESS*
58 PSYCHOPATHOLOGY*:ME
59 PSYCHOPATHOL*
60 AUTOGENIC-TRAINING*:ME
61 AUTOGENIC*
62 HEALTH-EDUCATION*:ME
63 [PATIENT near EDUCAT*]
64 [SELF near MANAGE*]
65 PSYCHOEDUCAT*
66 exp Rehabilitation/
67 exp Rehabilitation centers/
68 exp Rehabilitation nursing/
69 rehabil$.tw.
70 Aftercare/
71 aftercare.tw.
72 Convalescence/
73 Convalescen$.tw.
74 Recuperat$.tw.
75 Exp Heart diseases/rh [rehabilitation]
76 Patient education/
77 (patient adj2 educat$).tw.
78 counsel$.tw.
79 (behavi$ adj2 therap$).tw.
80 Psychosocial$.tw.
81 ((lifestyle or life-style) ADJ2 intervent$).tw.
82 ((lifestyle or life-style) ADJ5 (intervent$ or program$ or treatment$)).tw.
83 Nurse practitioners/
84 "nurse practitioner$".tw.
85 (secondary ADJ5 prevent$).tw.
86 Education/ or patient education/
87 Home care services/
88 Home.mp. or home nursing/
89 Self care/
90 Telephone/
91 Ambulatory care/
92 Cardiac rehabilitation.mp.
93 Heart manual.mp.

Figure 4.3 – 12
Final version of the search strategy link diagram