ORIGINAL ARTICLE

Discussing State-of-the-Art Spatial Visualization Techniques Applicable for the Epidemiological Surveillance Data on the Example of *Campylobacter* spp. in Raw Chicken Meat

C. Plaza-Rodríguez, B. Appel, A. Kaesbohrer and M. Filter

Federal Institute for Risk Assessment, Department Biological Safety, Unit Epidemiology, Zoonoses and Antimicrobial Resistance, Berlin, Germany

Impacts

- This study presents results of a review on the visualization techniques for epidemiological data including those for visualization of data uncertainty.
- A collection of GIS charts representing the prevalence of *Campylobacter* spp. in raw chicken meat obtained from the German Zoonoses Monitoring in 2011 is presented, including choropleth, cartogram, graduated symbol, dot-density, adjacent and coincident maps, which in part are capable of representing information on uncertainty.
- No single visualization technique outperforms in visualizing prevalence data or prevalence data together with the associated uncertainty. As a consequence, it is recommended to establish a dialogue between end-users and epidemiologists in order to determine which technique should be used in each case. This decision should consider previous knowledge and habits of end-user as well as the specific objective to be achieved with the visualization of data.

Keywords:
Spatial visualization techniques; spatial epidemiology; visualization of uncertainty; *Campylobacter*; geographic information systems

Correspondence:
C. Plaza-Rodríguez, Federal Institute for Risk Assessment, Department Biological Safety, Unit Epidemiology, Zoonoses and Antimicrobial Resistance, Marienfelde, *Diedersdorfer Weg* 1. D -12277 Berlin, Germany. Tel: +49 30 18412 2361; Fax: +49 30 18412 2952; E-mail: Carolina.Plaza-Rodriguez@bfr.bund.de

Received for publication December 1, 2014

doi: 10.1111/zph.12231

Summary

Within the European activities for the 'Monitoring and Collection of Information on Zoonoses', annually EFSA publishes a European report, including information related to the prevalence of *Campylobacter* spp. in Germany. Spatial epidemiology becomes here a fundamental tool for the generation of these reports, including the representation of prevalence as an essential element. Until now, choropleth maps are the default visualization technique applied in epidemiological monitoring and surveillance reports made by EFSA and German authorities. However, due to its limitations, it seems to be reasonable to explore alternative chart type. Four maps including choropleth, cartogram, graduated symbols and dot-density maps were created to visualize real-world sample data on the prevalence of *Campylobacter* spp. in raw chicken meat samples in Germany in 2011. In addition, adjacent and coincident maps were created to visualize also the associated uncertainty. As an outcome, we found that there is not a single data visualization technique that encompasses all the necessary features to visualize prevalence data alone or prevalence data together with their associated uncertainty. All the visualization techniques contemplated in this study demonstrated to have both advantages and disadvantages. To determine which visualization technique should be used for future reports, we recommend to create a dialogue between end-users and epidemiologists on the basis of sample data and charts. The final decision should also consider the knowledge and experience of end-users as well as the specific objective to be achieved with the charts.

Introduction

Foodborne diseases are a significant public health burden with important economic and social effects (Altekruse and Swerdlow, 1996; WHO, 2002, 2014; Kuchenmuller et al., 2013). Among the most important causes of foodborne infections, *Campylobacter* spp. is the most frequently reported causative agent in the EU (European Union) (EFSA, 2014b; EFSA and ECDC, 2014), being the fresh broiler meat the major source of human campylobacteriosis (Lin, 2009; EFSA and ECDC, 2014). In Germany, campylobacteriosis is the most common bacterial diarrhoeal disease in humans since 2007 (RKI, 2008). The Robert Koch Institute reported in 2011 a total of 71 307 cases, which corresponds to an incidence of 87.2 cases per 100 000 people (RKI, 2012).

To reduce the emergence of zoonotic infections, the EU has elaborated the 'European Union System for the Monitoring and Collection of Information on Zoonoses' based on Directive 2003/99/EC, intended to identify the sources of the most common pathogens causing foodborne and zoonotic diseases (EFSA, 2014a). In Germany, this is partly achieved by the National Zoonoses Monitoring, where all the data are collated and published in the National Zoonoses Report (Käsbohrer et al., 2009; BfR, 2013).

Epidemiological monitoring systems are an imperative instrument for protecting consumers from this health threat and maintaining the safety of the food supply (FAO, 2004). This monitoring generally includes a description of the geographical trends (Berkelman et al., 2009), revealing important clues that can help the decision-making (WHO, 2008). Spatial epidemiology becomes then a fundamental tool as it can describe, quantify and explain the geographical variations of diseases or contaminated items (Pina et al., 2010). In the last years, spatial epidemiology has experienced a great progress due to the advances in analytical methods such as GIS (Geographic Information Systems) and Spatial Analysis (Elliott and Wartenberg, 2004; Gomez-Rubio et al., 2004; Goodchild and Haining, 2005). GIS software solutions have evolved significantly in recent years and even when they are more complex, they are often easy-to-use products (Malczewski, 2004) with graphical user interfaces also supporting non-expert users. Currently, we can find hundreds of GIS tools. Steiniger and Bocher (2009) and Steiniger and Hunter (2013) carried out an overview on the current free and open source GIS software solutions. There are also many other commercial GIS tools used for epidemiologist purposes as ArcView ™, ArcGIS ™, MapInfo ™, Maptitude ™, Idrisi ™ and Geomedia ™ among others (Malczewski, 2004). ArcGIS is one of the most consolidated tools in this area. Moreover, it has a specific website with blog, forums, videos and an online help library allowing non-expert users to create their own maps (ESRI, 2015).

One of the most important areas within the spatial epidemiology is the mapping of the adverse events (Elliott and Wartenberg, 2004). Mapping has a long tradition and has been already used for many epidemiological purposes related to health (Elliott and Wartenberg, 2004; Rytkonen, 2004) in both human and animals (Norstrom, 2001). However, currently the use of GIS has greatly facilitated the mapping, allowing the storage, organization and processing of spatial data to be shown in the form of maps (Rytkonen, 2004; Goodchild and Haining, 2005; Madrid-Soto and Ortiz-López, 2005; Crampton, 2010; Rodrigues-Silveira, 2013).

When mapping is performed, one of the key points to be considered is the selection of the most suitable visualization technique. This election will determine the interpretability of the epidemiological data and therefore the decisions that will be taken based on this interpretation. EFSA (European Food Safety Authority) is the trendsetting institution in the identification of the most appropriate techniques for epidemiological spatial data visualization and their interpretation (EFSA, 2009). Which type of chart is chosen by EFSA strongly depends on the objectives of the respective study and on the nature of the data to be visualized. Choropleth maps have been widely used in public health and epidemiology (Cromley and Cromley, 2009); in the EU, it is used by default in the visualization process of epidemiological information (EFSA, 2009, 2011). For example, in the EFSA report on the prevalence of *Salmonella* in the European Union in 2008 (EFSA, 2008a,b), prevalence values of different *Salmonella* serovars were represented using simple choropleth maps. As can be easily seen from those choropleth maps, prevalence values of each serovar have been presented in different maps, instead of being included in just one map to facilitate comparisons. Moreover, uncertainties are not included; thus, correct interpretation of data is hampered. These and other limitations of choropleth maps might finally lead to misinterpretations (Dykes and Unwin, 1998; Cromley and Cromley, 2009), and it seems reasonable to explore alternative spatial display formats able to describe the epidemiological events in a more precise way.

Alternative chart types include cartograms, maps with graduated symbols or dot-density maps; some of them overcome the limitations of choropleth maps.

In addition, data used to create epidemiological maps are usually not exact, that is they always have some level of uncertainty. Therefore, it is important to visualize data together with its uncertainty, which would support decision-making if this is based on spatial data (Kardos et al., 2004). Several publications addressed this problem, by extending the classical visualization techniques such that it is possible to represent data together with uncertainty information (visualization of uncertainty) (MacEachren, 1992; Kardos et al., 2004; Viard et al., 2011).

The objective of this study was to explore suitable visualization techniques for epidemiological data including charts for visualization of data uncertainty.

Materials and Methods

Data

Data from the German Zoonoses Monitoring programme were used (Table 1). In the sampling plan, the sample size for each Federal State was proportional to the human population in the respective region (StatistischesBundesamt, 2014). The overall aim was to collect at least 385 samples to calculate the national prevalence with a precision of at least 5%.

The data set contains 430 results of raw chicken meat samples taken by the competent authorities in the 16 Federal States in Germany at retail level and analysed in the official regional investigation centres for the presence of *Campylobacter* spp. by ISO 10272-1:2006 method. Data were collected on national level by the Federal Office for Consumer Protection and Food Safety (BVL) and transferred to the Federal Institute for Risk Assessment for further analysis. Prevalence was estimated for each Federal State according to the following equation:

$$P = X/n$$

where P is the prevalence of *Campylobacter* spp., X is the number of positive samples, and n is the number of total tested samples in the corresponding state. Prevalence expressed as a percentage was used to create the dot-density map.

To estimate the uncertainty associated with the prevalence, we used the half of the confidence interval (1.96*SE), by following the approach proposed by Agresti and Coull (1998) for constructing 95% confidence intervals of binomial proportions:

$$\mathrm{CI} = \tilde{p} \pm 1.96 * \mathrm{SE}$$

$$\mathrm{SE} = \sqrt{\tilde{p} * (1 - \tilde{p})/(n + 1.96^2)}$$

$$\tilde{p} = (X + 1.96^2/2)/(n + 1.96^2)$$

where CI is the 95% confidence interval, SE is the standard error, and $\tilde{p}$ is the corrected prevalence according to Agresti and Coull (1998). All data are provided in Table 1.

Maps

The following visualization techniques were compared:

Table 1. Data from the Zoonoses Monitoring completed in Germany on the prevalence of *Campylobacter* spp. in raw chicken meat collected at retail (2011)

German Federal States Information						Zoonoses Monitoring-2011 Results						
Code	Name	RS	Area (km²)	Inhabitants (Number)	Density (inhabitants/km²)	No Samples[a]	Positive Samples	P	P (%)	$\tilde{p}$	SE	1.96*SE
BB	Brandenburg	12	29 654	2 449 193	83	12	0	0.000	0.00	0.121	0.082	0.161
BE	Berlin	11	892	3 421 829	3838	28	4	0.143	14.29	0.186	0.069	0.135
BW	Baden-Württemberg	8	35 751	10 631 278	297	52	22	0.423	42.31	0.428	0.066	0.130
BY	Bayern	9	70 550	12 604 244	179	80	28	0.350	35.00	0.357	0.052	0.103
HB	Bremen	4	419	657 391	1568	3	1	0.333	33.33	0.427	0.189	0.371
HE	Hessen	6	21 115	6 045 425	286	14	2	0.143	14.29	0.220	0.098	0.192
HH	Hamburg	2	755	1 746 342	2312	8	4	0.500	50.00	0.500	0.145	0.285
MV	Mecklenburg-Vorpommern	13	23 212	1 596 505	69	8	4	0.500	50.00	0.500	0.145	0.285
NI	Niedersachsen	3	47 614	7 790 559	164	34	6	0.176	17.65	0.209	0.066	0.130
NW	Nordrhein-Westfalen	5	34 110	17 571 856	515	90	32	0.356	35.56	0.361	0.050	0.097
RP	Rheinland-Pfalz	7	19 854	3 994 366	201	30	6	0.200	20.00	0.234	0.073	0.143
SH	Schleswig-Holstein	1	15 800	2 815 955	178	11	1	0.091	9.09	0.197	0.103	0.202
SL	Saarland	10	2569	990 718	386	4	2	0.500	50.00	0.500	0.179	0.350
SN	Sachsen	14	18 420	4 046 385	220	23	12	0.522	52.17	0.519	0.096	0.189
ST	Sachsen-Anhalt	15	20 452	2 244 577	110	21	6	0.286	28.57	0.319	0.094	0.183
TH	Thüringen	16	16 173	2 160 840	134	12	10	0.833	83.33	0.752	0.108	0.213
DE	**Deutschland**		**357 340**	**80 767 463**	**226**	**430**	**140**	**0.335**	**33.474**	**0.364**	**0.101**	**0.198**

RS: Regional code; P: prevalence; P (%): prevalence expressed as a percentage; $\tilde{p}$: corrected prevalence according to Agresti and Coull, 1998; SE: standard error according to Agresti and Coull (1998); 1.96*SE: half of the 95% confidence interval according to Agresti and Coull (1998).

[a]The sample size calculation was performed on the national level and was allocated to the regions proportional to the population, so uncertainty is influenced by this fact.

Global values from Germany have been highlighted in bold.

1. *Choropleth map:* A choropleth or *area-value map* displays the measured data in connection with the administrative boundaries (Slocum et al., 2005). Data are grouped into two or more bins that are coloured using different colours to illustrate the spatial differences in the measured magnitude (Brewer and Pickle, 2002).
2. *Cartogram: A* cartogram or *value-by-area map* changes the size of the spatial or enumeration units depending on the value of the attribute to be visualized (Keim et al., 2002). There are different types of cartograms, but all of them present some level of shape and/or topology distortion (Dent, 1996; Slocum et al., 2005), mainly depending on whether they have to maintain the connectivity with their adjacent enumeration units or not, and whether they replace the enumeration units by a geometrical figure with a size depending on the variable (Dorling, 1996; Kreveld and Speckmann, 2007).
3. *Graduated and proportional symbol maps:* Graduated and proportional symbol maps use symbols of different sizes to represent the variable of interest (Brewer and Campbell, 1998). In proportional symbol maps, data are unclassed and the symbols are proportional to the numerical values of these data (Gruver and Dutton, 2014). In this study, we present a graduated symbol map in which data are classified into different classes that are then correlated with a given size of the symbol.
4. *Dot-density map:* Dot-density or *density point map* uses dots or points placed on a map. Each dot represents a specific number of epidemiological events (Lavin, 1986). Unlike the graduated symbol maps, all points are of the same size. Dot-density map provides very good visual impression of the relative density with which the data are presented in space (Berg et al., 2004).

Several approaches for visualizing uncertainty have been proposed. This include (i) adjacent representations, (ii) coincident representations and (iii) interactive techniques (Kardos et al., 2004). In this study, adjacent maps and coincident maps were applied to visualize prevalence together with its associated uncertainty. Interactive techniques were not included as they are not applicable in written reports.

1. *Adjacent maps:* The adjacent map consists on the representation of two maps, next to each other, one for primary data and the other for the uncertainty associated with these data (MacEachren, 1992; MacEachren et al., 1998; Viard et al., 2011).
2. *Coincident map*: A coincident map can be considered to be a special case of bivariate mapping (MacEachren et al., 1998), as uncertainty is integrated into the same map as the primary data. In this type of map, the uncertainty is integrated by changing the colour characteristics or by overlapping a new layer with different symbols or textures (MacEachren et al., 1998; Viard et al., 2011).
   a. *Combination of texture and colour map:* The combination of texture and colour map is a coincident technique that uses a colour fill to represent primary data and an overlapping layer with different textures to represent the different levels of uncertainty (MacEachren et al., 1998).
   b. *Value-by-alpha map:* The value-by-alpha map is a visualization technique in which two variables of known relation are mapped (Roth et al., 2010). It can be used as a coincident map to visualize data and the uncertainty associated, because the latter is encoded in the same map by modifying the alpha channel of the colours from primary data map (Woodruff, 2010).

To facilitate a 'fair' comparison between the different visualization techniques, some visualization parameters were fixed for all generated maps. (i) Most authors agree that 5 –7 classes are appropriate (Gilmartin, 1981; Brewer and Pickle, 2002; Roth et al., 2010) to find a balance between data generalization (few classes) and readability (more classes are harder to be properly interpreted) (Gilmartin and Shelton, 1990; Roth et al., 2010). In our case, data were classified into five bins. Zero prevalence, whenever reported, was included within the lower prevalence category. For uncertainty visualization maps, the calculated uncertainty measure was grouped into three bins, as it is recommended by Roth et al. (2010). (ii) Different opinions can be found about which method should be used to divide the data into categories (Jenks, 1967; Jenks and Caspall, 1971; Brewer and Pickle, 2002), as the change in the classification method can change also how the map looks and the message it sends (EFSA, 2009). Two of the most common classification methods are natural breaks (where classes are defined aiming to minimize within-class variance and maximize between-class variance) (Brewer and Pickle, 2002) and quantile breaks (where data are divided into pre-defined numbers of classes which contain the same number of events) (EFSA, 2009). For choropleth, cartogram, graduated symbol and uncertainty visualization maps, prevalence data were grouped using cut-points generated using natural breaks. (iii) It has been also demonstrated that hue affects the accuracy rates and reaction times when interpreting a map (Gilmartin and Shelton, 1990). As our maps aim to show not qualitative but quantitative data, a scheme of sequential colours with a single-hued was used for choropleth map, cartogram and choropleth maps used as a basemap to display uncertainty, following the criteria of darker corresponding to a higher prevalence (McGranaghan, 1993). For the combination of colour and texture map, a new layer with different textures was overlaid to a choropleth map. Sequential colour multihued was used for value-by-alpha maps, overlaying on a neutral colour background (white and black).

GIS software

ESRI's ArcGIS 10.2.2 for Desktop (ESRI, 2014) was used to create the maps in this study. In addition, the open source software ScapeToad (ChôrosLaboratory, 2014) was used to create the cartogram. This software uses the Gastner and Newman (2004) method to density-equalize maps, by transforming the size and sometimes the shape of enumeration units. Output values generated by ScapeToad were then converted into a shapefile and imported into ArcGIS to visualize the cartogram.

The intensity of the colour was automatically assigned by the software tool. No changes were made in the intensity value, as we consider that the predefined values allowed a correct interpretation of the data.

For value-by-alpha maps, multihued colours were varied in transparency depending on the uncertainty by modifying their alpha channel from 10–20% (for the highest uncertainty class) to 100% (for the lowest uncertainty class).

Results and Discussion

Prevalence visualization techniques

A general known issue with maps discussed in this study (except the dot-density maps) is that they give the impression that all enumeration units belonging to a specific category (with the same colour hue or symbol size) have exactly the same value and that this value change abruptly at the boundaries. As illustrated in the following case study, this impression might be misleading.

Choropleth map

Figure 1a shows the choropleth map created for the *Campylobacter* spp. prevalence data in the Federal States. From this map, it is not difficult to extract information related to prevalence as the different blue hues can be effortlessly differentiated, and darker blue hues correspond to a higher prevalence in distinct Federal States.

Choropleth maps are easy to create in most of the GIS software and are easily understandable by the map readers (Gruver and Dutton, 2014). Moreover, this type of map allows interpreting several variables at the same time, for example by including some other symbols over the basemap (bivariate map).

However, despite its popularity, its limitations can lead to an incorrect data interpretation (Roth et al., 2010). Two main limitations have been extensively described:

1. *Enumeration Units Size or Small number problem:* Size of enumeration units can vary greatly. For example, in Germany, the size of the Federal States and population living in the regions is highly variable, as shown in Table 1. When raw data are represented, larger enumeration units would dominate the perception of the map, overstating in our case the magnitude of the of positive samples (Cromley and McLafferty, 2002). Therefore, the number of positive samples was standardized by the total number of samples taken in each Federal State.
2. *Modifiable Areal Unit Problem (MAUP):* Interpretation of a choropleth map depends on the boundaries of the enumeration units (Openshaw, 1984; Heywood et al., 1988). For example, data aggregated by countries might not be able to display relevant regional differences compared to data aggregated by regions (EFSA, 2009).

Despite these limitations, we appraise that, in case of prevalence data, choropleth maps work reasonably well.

Cartogram

Figure 1b shows the generated cartogram. In this map, the size of the Federal States was modified depending on the prevalence value so that the shape and topology of the original geography were distorted. Moreover, to facilitate the map interpretation, the prevalence values were also discretized into five classes (same as for Fig. 1a) and then assigned to the area using the choropleth technique overlaying the cartogram. As can be observed from Fig. 1b, areas with a higher prevalence are clearly highlighted. Those Federal States that have originally small areas but high prevalence values are now displayed as larger areas. On the other side, big areas of the original map became nearly invisible in case where the prevalence is low.

This demonstrates that a cartogram is a suitable technique for visualization of prevalence information, and it is able to attract the attention to spatial units with high prevalence rates which might have been overseen otherwise. However, those regions with low prevalence rates, like Brandenburg, might now been overseen, which is not desirable. Furthermore, it has been described that in some cases, the changes in topology and shape are overly large and map readers have problems to recognize the original enumeration units (Roth et al., 2010). In our example, the region Brandenburg is hardly recognizable anymore, and also the fact that the Federal State Bremen consists of two cities surrounded each by the Federal State Niedersachsen creates a quite disturbing picture. In addition, when two enumeration units have the same size but different shapes, map reader assumes that they have different sizes (Indiemapper, 2010). Here, we solved this issue colour-coding the data in parallel.

Cartograms have gained popularity as they overcome issues related to choropleth maps (Roth et al., 2010). This type of map can also be used to encode two variables (bivariate) by adding choropleth-like fills to each enumeration unit in relation to the second variable, or using it as a basemap for other type of maps such as density point maps or graduated symbol map.

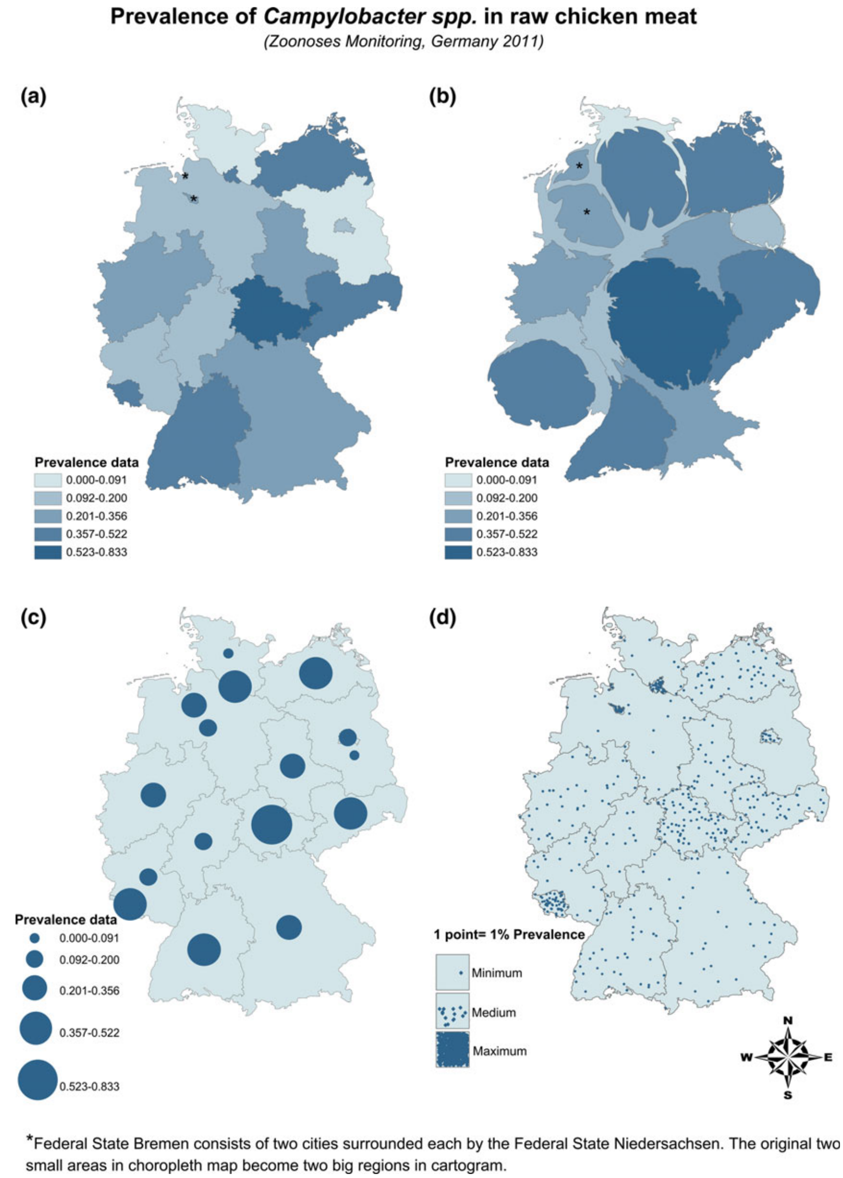


Fig. 1. Prevalence visualization maps: Maps depicting the prevalence of *Campylobacter* spp. in raw chicken meat (National Zoonoses German Report, 2011). (a) Choropleth map, (b) cartogram, (c) graduated symbol map and (d) dot-density map.

Graduated symbols map

Figure 1c shows the created graduated symbol map. In this case, graduated circles of different sizes were created depending on the prevalence value where larger circles correspond to higher prevalence. The differences in circle sizes were adjusted such that it allows effective and correct interpretation of the prevalence data in the map. As can be observed in the figure, this type of map overcomes the enumeration unit problem of choropleth maps, as the size of the symbol depends just on the prevalence and not on the size of enumeration unit (Brewer and Campbell, 1998). Small enumeration units with high prevalence values can also have a large symbol over them.

Graduated symbol maps are very flexible. They can be developed from raw data and standardized data, they allow to display several variables using compound symbols (Brewer and Campbell, 1998; Nelson, 2000), and they can be used for data attached to a precise location or data attached to geographic areas as is in our case (Brewer and Campbell, 1998).

A known issue is the so-called Ebbinghaus illusion which can provoke that two identical circles appear to be of different sizes depending on the contour that surround them (Gilmartin, 1981). Other issues related to graduated symbol map can easily been overcome:

1. By chosen carefully the size of the symbols, it is possible to avoid symbol overlapping which hinders the correct

interpretation of the data (Groop and Cole, 1978; Gruver and Dutton, 2014).
2. By classifying data into few classes, differences between symbol size will be easily noticeable (Brewer and Campbell, 1998), allowing map readers to estimate variable values properly.

However, as we can see from our graduated symbol map, Federal States with small areas can be hidden completely by the symbol that depicts its prevalence, which could hamper the interpretation of results.

Dot-density map/density point map

Figure 1d shows the dot-density map. Prevalence values were expressed in unit [%] and were depicted in this map by the use of little points that were randomly distributed over the corresponding Federal State. The size and value of the dots were carefully selected to achieve the best interpretation of the data. As is recommended by Olson (1975) and Gruver and Dutton (2014), the legend includes the value of each dot (one dot= 1% prevalence) and also an example of the appearance that low, medium and high density of dots have on the map. As we can see from Figure 1d, it is difficult to extract numerical prevalence values from this type of map. However, Federal States with higher prevalence can be distinguished easily by the presence of a higher density of points. Nevertheless, even when some authors argue that the visual impression of the density provided by this type of map is really useful to interpret the data (Lavin, 1986), one has to be aware that the density impression still depends on the size of the enumeration unit, as, for example, small ones like Bremen or Hamburg are more noticeable than other enumeration units that have the same number of dots but higher area.

Dot-density maps have become a popular technique to visualize density distributions (Berg et al., 2004), because although in our case our data have been aggregated by enumeration units, this is not a prerequisite for this type of maps, and hence, this is not a limitation to this visualization technique. It also allows more than one attribute to be represented using dots of different colours encoding different variables. However, we think that this map is not the best option to represent prevalence, but rather to represent the total number of investigated and positive samples at each Federal State using two dot types.

Combined prevalence and uncertainty visualization

A specific challenge is the graphical representation of uncertainty associated with prevalence data. In this study, we aim at providing examples on applicable visualization techniques applicable for this kind of data.

Although it has been initially anticipated that graphical representation of uncertainty together with the data in a map will disturb the reader (McGranaghan, 1993), many studies have demonstrated that inclusion of uncertainty could clarify the interpretation of results (Leitner and Buttenfield, 2000; Aerts et al., 2003; Viard et al., 2011). Particularly, this applies when data are used in the decision-making process (Brodlie et al., 2012) and when variation in results may have a great transcendence (Hengl et al., 2002; Aerts et al., 2003), such as in the public health domain. Nonetheless, most epidemiological maps generated to date do not incorporate uncertainty information. For correct interpretation of uncertainty values of the selected database, it is important to highlight that the uncertainty values were not correlated with the size of the enumeration unit.

In our example, uncertainty is mainly affected by the sample size, which was proportional to the population size, of the different Federal States.

Adjacent maps

Figure 2a shows two separate choropleth maps representing on the left side the prevalence of *Campylobacter* spp. and on the right side the uncertainty related to this prevalence. These maps simplify the visualization of each aspect for non-experts (Gerharz et al., 2012) but have to be mentally overlaid to determine the uncertainty associated with the prevalence of a specific Federal State (Gruver and Dutton, 2014). Each map has its own legend, defining the different levels of prevalence and uncertainty, respectively.

Even when this map avoids the visual overload of coincident maps, it can be clearly seen from Fig. 2a that it is hard to connect mentally information from both maps, as it requires a compound comparison in multiple areas (Harrower, 2003).

Coincident maps

Combination of texture and colour map

Figure 2b shows an example of a coincident representation depicting uncertainty by the combination of both texture and colour in the very same map. A choropleth map was used to represent the prevalence values. Over this, a new layer with different textures was placed to represent the associated uncertainty values. Figure 2b demonstrates that it is possible to combine different textures with colour hue to depict prevalence and uncertainty in an effective way. This is in line with observations made by other authors (MacEachren et al., 1998). With coincident maps, users cannot ignore the uncertainty information (Evans, 1997; Edwards and Nelson, 2001; Viard et al., 2011), but the perception depends on their expertise (Tversky and

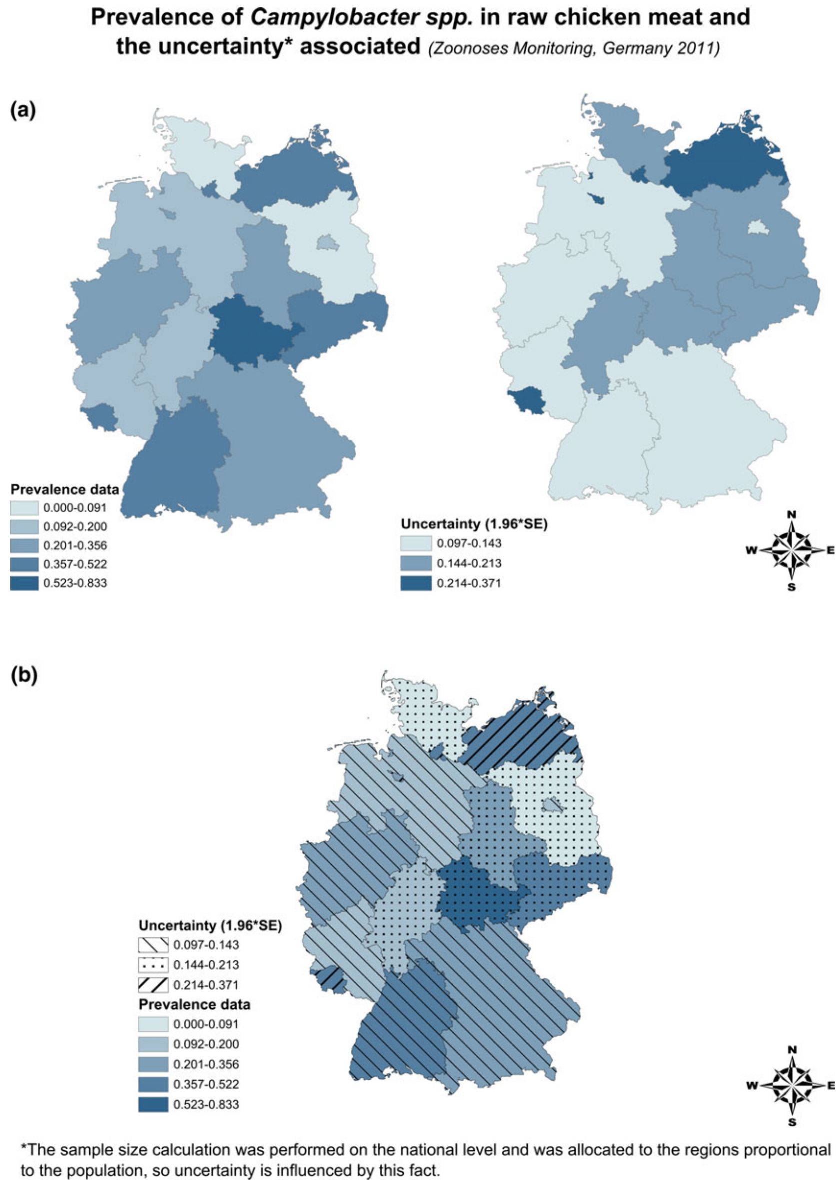


Fig. 2. Uncertainty visualization maps I: (a) Adjacent maps and (b) combination of colour and texture map, representing the prevalence of *Campylobacter* spp. in raw chicken meat and its uncertainty (*1.96*Standard Error*). (National Zoonoses German Report, 2011).

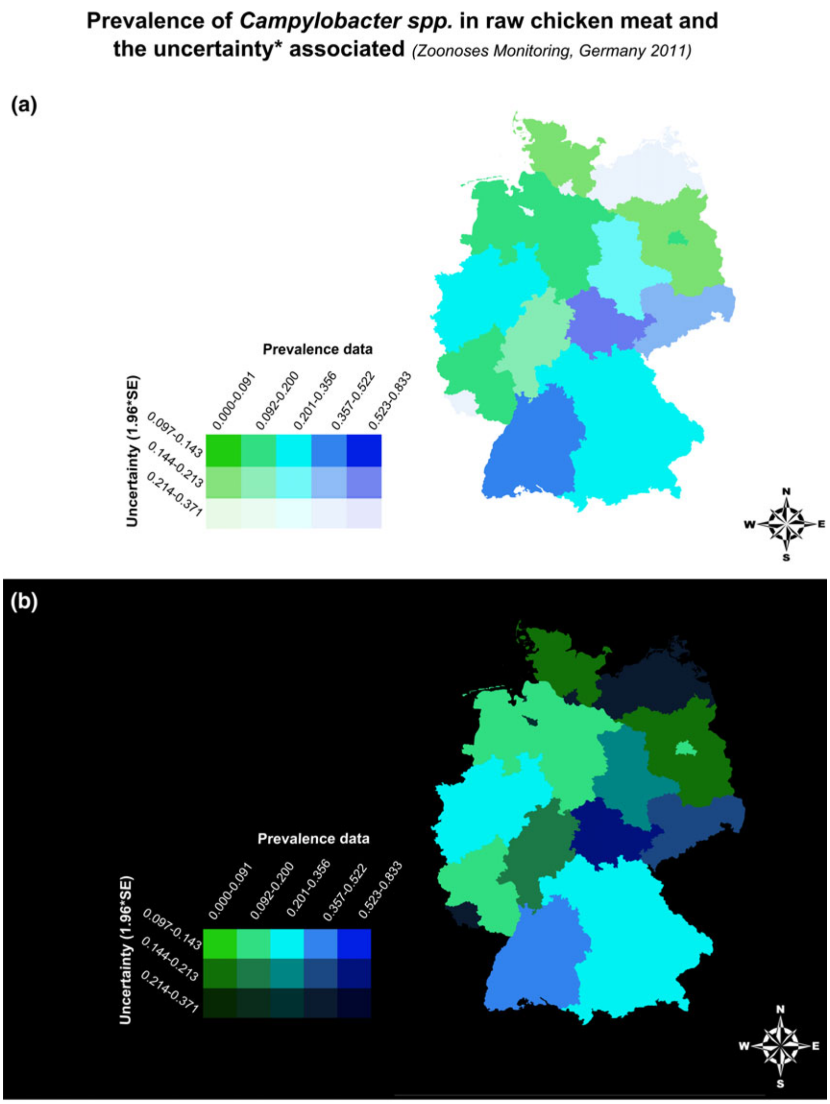
Kahneman, 1974; Gerharz and Pebesma, 2009). Using this visualization technique, one can interpret both parameters at a glance, as the textures from the upper layer do not interfere with interpretation of the underlying choropleth map.

Value-by-alpha map

Figure 3 provides two versions of a value-by-alpha map. As it is described by Roth et al. (2010), three components had to be considered in the design of these maps: (i) the *variable of interest* (prevalence of *Campylobacter* spp.), (ii) the *equalizing variable* (uncertainty) that is symbolized by the alpha value and equalizes the map and (iii) the *modifying colour* or background colour (white for Fig. 3a and black for Fig. 3b). This colour modifies the original colour for prevalence, as its alpha value changes depending on the level of uncertainty.

As a result, Federal States with a high uncertainty are drawn more transparent, nearly invisible, hinting the background colour (black or white) and masking the colour hue of the variable of interest (Roth et al., 2010; Gruver and Dutton, 2014). On the other hand, Federal States with low uncertainty are fully opaque and stand out over the rest. The legend was created by collecting all the possible combinations of colour and transparency corresponding to the combinations of prevalence and the uncertainty.

Both maps created with different background colours prove to be useful for our purpose. However, as other authors have argued, it seems that black background



*The sample size calculation was performed on the national level and was allocated to the regions proportional to the population, so uncertainty is influenced by this fact.

Fig. 3. Uncertainty visualization maps II: Value-by-alpha maps representing the prevalence of *Campylobacter* spp. in raw chicken meat and its uncertainty (*1.96*Standard Error*) (National Zoonoses German Report, 2011). (a) Value-by-alpha map with white background and (b) value-by-alpha map with black background.

perform better than the white one, because it is able to stand out areas of low uncertainty more clearly (Roth et al., 2010).

Even when value-by-alpha maps are suitable to display both prevalence and uncertainty, as we can see from Fig. 3a and b, they are difficult to interpret for non-experts as colour variation depends on the combination of the three components that we have described previously and the general rule that ‘darker equals more’ is not applicable (Schweizer and Goodchild, 1992; Harrower, 2003).

Summary and recommendations

Nowadays visual representations have become indispensable tools for epidemiological data presentation and decision-making. This comprise not only archiving and communication of results, but also its analysis and exploration, helping to identify trends in the emergence of new foodborne diseases as, for example campylobacteriosis. This leads to the need to optimize the use and design of maps in order to support decision-making by public health policies. Large discussions about spatial visualization techniques can be found in the scientific literature. Most scientific papers do not opt 100% for one visualization technique against the others, but rather they describe advantages and disadvantages and even the percentage of efficiency that a particular map has over the rest when they are interpreted by a population of interest. However, many contradictory opinions can be found between authors. There is a common theme among them criticizing the application of choropleth maps, due to limitations mentioned above. This is also why many new visualization techniques are proposed nowadays. As we can take out from our comparison, choropleth map is the simplest way to depict the prevalence of *Campylobacter* spp. However, for deeper studies, more sophisticated maps as cartogram, graduated symbol and dot-density maps can be more suitable. If we want to draw attention to areas of highest prevalence, we can use a cartogram as long as stakeholders are used to interpret this type of representation. For inexperienced users, as, for example, visitors of public websites looking for general information on foodborne diseases, we recommend to use a graduated symbol or dot-density maps, which would highlight the magnitude of the variable without distorting the shape of the map. If the goal is to present two different prevalences on the same map in order to compare them, we recommend to combine a choropleth map and a cartogram. However, these types of maps are most likely not appropriate for non-expert users.

With respect to maps displaying uncertainty levels together with primary data, there is no clear sentiment in the literature whether adjacent or coincident representations perform better. From our case study, we conclude that the combination of colour and texture may be the most effective coincident technique for unexperienced users, as it is able to evince uncertainty without hindering the detection of patterns in the primary data (MacEachren et al., 1998). New techniques that combine colour and textures such as the rhombus trustree tessellation map might also be useful visualization techniques (Kardos et al., 2004).

In summary, we have to realize that, as Stewart and Kennelly (2010) said, ‘the utility of a map to the map user depends on a number of factors such as geographic complexity of the phenomenon to be mapped, the decision to present data as either classed or unclassed, the method by which data will be symbolized, and the ability of the map user to interpret the resulting map’. Therefore, there is not a map that encompasses all the necessities so that there is not a ‘perfect’ visualization technique to be recommended. Instead an interactive dialogue between end-users and epidemiologist has to be established in order to identify which technique should be used in the specific setting. This selection process should also consider the professional background and experiences of map readers and balance this with the objectives to be achieved with the map (Gerharz and Pebesma, 2009). It is recommended to test the usability of different visualization techniques beforehand to identify the most effective visualization technique. The data set and chart collection from this study can serve as a basis for this type of discussion. In addition, it is also necessary to judge carefully on the variables to be represented, in order to avoid the cluttering of information (Leitner and Buttenfield, 2000; Viard et al., 2011). In the public health domain, it could be useful to combine simple maps with more sophisticated ones, in order to get a general view and a deeper assessment of the epidemiological data (Gerharz and Pebesma, 2009).

Acknowledgements

This work is part of the project PROMISE (Project no. 265877) funded by the 7th Framework Programme of the European Union FP7, Knowledge-based Bio-Economy (KBBE). We thank Dr. Beutlich for her previous report and Dr. Correira-Carreira for the thorough review of the manuscript and his useful suggestions.

References

Directive 2003/99/EC of the European Parliament and of the Council of 17 November 2003 on the monitoring of zoonoses and zoonotic agents, amending Council Decision 90/424/EEC and repealing Council Directive 92/117/EEC. *Off. J. Eur. Union L* 325, 31.

Aerts, J. C. J. H., K. C. Clarke, and A. D. Keuper, 2003: Testing popular visualization techniques for representing model uncertainty. *Cartogr. Geogr. Inf. Sci.* 30, 249–261.

Agresti, A., and B. A. Coull, 1998: Approximate is better than ‘exact’ for interval estimation of binomial proportions. *Am. Stat.* 52, 119–126.

Altekruse, S. F., and D. L. Swerdlow, 1996: The changing epidemiology of foodborne diseases. *Am. J. Med. Sci.* 311, 23–29.

Berg, M., P. Bose, O. Cheong, and P. Morin, 2004: On simplifying dot maps. *Comp. Geom. Theor. Appl.* 27, 43–62.

Berkelman, R. L., P. S. Sullivan, and J. W. Buehler, 2009: Public health surveillance. In: Detels, R., R. Beaglehole, M. A. Lansing, and M. Gulliford (eds), Oxford Textbook of Public Health. Oxford University Press, Oxford.

BfR, 2013: Erreger von Zoonosen in Deutschland im Jahre 2011. In: Hartung, M., and A. Käsbohrer(eds), BfR Wissenschaft 05/2013, pp. 1–268. Bundesinstitut für Risikobewertung, Berlin, Germany.

Brewer, C. A., and A. J. Campbell, 1998: Beyond graduated circles: varied point symbols for representing quantitative data on maps. *Cartograph. Perspect.* 29, 6–25.

Brewer, C. A., and L. Pickle, 2002: Evaluation of method for classifying epidemiological data on choropleth maps in series. *Ann. Assoc. Am. Geogr.* 92, 662–681.

Brodlie, K., R. A. Osorio, and A. Lopes, 2012: A review of Uncertainty in Data Visualization. In: Dill, J., R. Earnshaw, D. Kasik, J. Vince, and P. Chung-Wong (eds), Expanding the Frontiers of Visual Analytics and Visualization. Springer-Verlag, London.

ChôrosLaboratory, 2014: ScapeToad v. 1.1. Available at: http://scapetoad.choros.ch/download.php(accessed on 10 August 2014).

Crampton, J. W., 2010: Mapping: A Critical Introduction to Cartography and GIS. Wiley-Blackwell, Oxford.

Cromley, R. G., and E. K. Cromley, 2009: Choropleth map legend design for visualizing community health disparities. *Int. J. Health Geogr.* 8, 1–11.

Cromley, E. K., and S. McLafferty, 2002: GIS and Public Health. Guilford Press, New York.

Dent, B. D., 1996: Cartography: Thematic map design. WCB/ McGraw-Hill, Boston, MA.

Dorling, D., 1996: Area Cartograms: Their Use and Creation, Concepts and Techniques in Modern Geography. Environmental Publications, Norwich.

Dykes, J., and D. Unwin, 1998: Maps of the Census: a rough guide. Available at: http://www.agocg.ac.uk/reports/visual/casestud/dykes/dykes.pdf(accessed on 14 October 2014).

Edwards, L., and E. Nelson, 2001: Visualizing data certainty: a case study using graduated circle maps. *Cartograph. Perspect.* 38, 19–36.

EFSA, 2008a: Report of the Task Force on Zoonoses Data Collection on the Analysis of the baseline survey on the prevalence of *Salmonella* in turkey flocks, in the EU, 2006-2007. *EFSA J.* 134, 1–91.

EFSA, 2008b: Scientific Report: analysis of the baseline survey on the prevalence of *Salmonella* in turkey flocks, in the EU, 2006-2007. Part B: factors related to *Salmonella* flock prevalence and distribution of *Salmonella* serovars. *EFSA J.* 198, 1–224.

EFSA, 2009: Statistical analysis of temporal and spatial trends of zoonotic agents in animals and food. Part I: critical review of the statistical analysis carried out on the Community Summary Report 2006 data. *EFSA J.* 253, 1–77.

EFSA, 2011: Statistical analysis of temporal and spatial trends of zoonotic agents in animals and food Part II: applications of spatial analysis and further developments of temporal analysis. *EFSA J.* 9, 2331.

EFSA, 2014a: Data dictionaries - guidelines for reporting data on zoonoses, antimicrobial resistance and food-borne outbreaks using the EFSA data models for the Data Collection Framework (DCF) to be used in 2014, for 2013 data. EFSA supporting publication 2014: EN-577, 102.

EFSA, 2014b: Food-borne zoonotic diseases: *Campylobacter*. Available at: http://www.efsa.europa.eu/en/topics/topic/campylobacter.htm(accessed on 5 August 2014).

EFSA and ECDC, 2014: The european union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2012. *EFSA J.* 12, 3547.

Elliott, P., and D. Wartenberg, 2004: Spatial epidemiology: current approaches and future challenges. *Environ. Health Perspect.* 112, 998–1006.

ESRI, 2014: ArcGIS for Desktop 10.2.2 ed.

ESRI, 2015: ArcGIS Resources. Available at: http://resources.arcgis.com/en/home/ (accessed on 18 June 2015).

Evans, B., 1997: Dynamic display of spatial data-reliability: does it benefit the map user? *Comput. Geoscic.* 23, 409–422.

FAO, 2004: Building Effective Food Safety Systems. Second FAO/WHO Global Forum of Food Safety Regulations. Building Effective Food Safety Systems. Food and Agriculture Organization of the United Nations, Bangkok, Thailand.

Gastner, M. T., and M. Newman, 2004: Diffusion-based method for producing density equalizing maps. In: Goodchild, M. F., (ed.), Proceedings of the National Academy of Sciences of the United States of America, pp. 7499–7504. National Academy of Sciences, Washington, DC.

Gerharz, L. E., and E. J. Pebesma, 2009: Usability of Interactive and Non-Interactive Visualisation of Uncertain Geospatial Information. In: Reinhardt, W., A. Krüger, and M. Ehlers, (eds.), Geoinformatik 2009 Konferenzband, pp. 223–230. ifgiPrints, University of Münster, Münster.

Gerharz, L. E., C. Autermann, H. Hopmann, C. Stasch, and E. Pebesma, 2012: Uncertainty visualisation in the ModelWeb. 9th EGU General Assembly 2012. Geophysical Research Abstracts. Vienna, Austria.

Gilmartin, P. P., 1981: Influences of Map Context on Circle Perception. *Ann. Assoc. Am. Geogr.* 71, 253–258.

Gilmartin, P., and E. Shelton, 1990: Choropleth maps on high resolution CRTs: the effects of number of classes and hue on communication. *Cartographica* 26, 40–52.

Gomez-Rubio, V., A. López-Quílez, and F. Verderjo, 2004: AEGIS: Herramienta de análisis epidemiológico en un sistema de información geográfica. *Congreso conjunto de epidemiología y biometría. II Jornada Científica SEE-SEB. "In Memoriam Juan Ferrándiz". Análisis espacial y temporal*. Valencia, Spain.

Goodchild, M. F., and R. P. Haining, 2005: SIG y análisis espacial de datos: perspectivas convergentes. *Investigaciones Regionales* 6, 175–201.

Groop, R. E., and D. Cole, 1978: Overlapping graduated circles/ magnitude estimation and method of portrayal. *The Canadian Cartographer* 15, 114–122.

Gruver, A., and J. A. Dutton, 2014: Geography 486. Cartography and Visualization. e-Education Institute. College of Earth and Mineral Sciences, The Pennsylvania State University.

Harrower, M., 2003: Representing uncertainty: Does it help people make better decisions? UCGIS Workshop: Geospatial Visualization and Knowledge Discovery Workshop, National Conference Center, Landsdowne, VA.

Hengl, T., D. J. J. Walvoort, and A. Brown, 2002: Pixel and colour mixture: GIS techniques for visualization of fuzziness and uncertainty of natural resource inventories. In: Hunter, G. J., and K. Lowell (eds), Proceedings of the 5th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Science (Accuracy 2002), 300–309. Department of Geomatics, University of Melbourne, Melbourne, Australia.

Heywood, I., S. Cornelius, and S. Carver, 1988: An Introduction to Geographical Information Systems. Addison Wesley Longman, New York.

Indiemapper, 2010. Available at: http://indiemapper.com/app/learnmore.php?l=index(accessed on 23 October 2014).

Jenks, G., 1967: The data model concept in statistical mapping. *Int. Yearbook of Cartography* 7, 186–190.

Jenks, G. F., and F. C. Caspall, 1971: Error on choropleth maps: definition, measurement, and reduction. *Ann. Assoc. Am. Geogr.* 61, 217–244.

Kardos, J., A. Moore, and G. Benwell, 2004: The Trustree for the Visualisation of Attribute and Spatial Uncertainty: usability Assessments. In: Whigham, P. A., and B. R. McLennan (eds), SIRC 2004: a Spatio-Temporal Workshop - Proceedings of the 16th Annual Colloquium of the Spatial Information Research Centre. University of Otago, Dunedin, New Zealand.

Käsbohrer, A., C. Wegeler, and B.-A. Tenhagen, 2009: EU-weite und nationale Monitoringprogramme zu Zoonoseerregern in Deutschland. *J. Verbr. Lebensm.* 4, 41–45.

Keim, D. A., S. C. North, C. Panse, and J. Schneidewind, 2002: Efficient Cartogram Generation: A Comparison. InfoVis 2002, IEEE Symposium on Information Visualization. Boston, Massachusetts.

Kreveld, M., and B. Speckmann, 2007: On rectangular cartograms. *Comp. Geom. Theor. Appl.* 37, 175–187.

Kuchenmuller, T., B. Abela-Ridder, T. Corrigan, and A. Tritscher, 2013: World Health Organization initiative to estimate the global burden of foodborne diseases. *Rev. Sci. Tech.* 32, 459–467.

Lavin, S., 1986: Mapping continuous geographical distributions using dot-density shading. *Am. Cartographer* 13, 140–150.

Leitner, M., and B. Buttenfield, 2000: Guidelines for the display of attribute certainty. *Cartogr. Geogr. Inf. Sci.* 27, 3–14.

Lin, J., 2009: Novel approaches for Campylobacter control in poultry. *Foodborne Pathog. Dis.* 6, 755–765.

MacEachren, A. M., 1992: Visualizing uncertain information. *Cartograph. Perspect.* 13, 10–19.

MacEachren, A. M., C. A. Brewer, and L. W. Pickle, 1998: Visualizing georeferenced data: representing reliability of health statistics. *Environ. Plann. A.* 30, 1547–1561.

Madrid-Soto, A., and L. M. Ortiz-López, 2005: Análisis espacial. In: Madrid-Soto, A., and L. M. Ortiz-López (eds), Análisis y síntesis en cartografía: algunos procedimientos. Universidad Nacional de Colombia, Bogotá.

Malczewski, J., 2004: GIS-based land-use suitability analysis: a critical overview. *Prog. Plann.* 62, 3–65.

McGranaghan, M., 1993: A cartographic view of spatial data quality. *Cartographica* 30, 8–19.

Nelson, E. S., 2000: Designing effective bivariate symbols: the influence of perceptual grouping processes. *Cartogr. Geogr. Inf. Sci.* 27, 261–278.

Norstrom, M., 2001: Geographical Information System (GIS) as a tool in surveillance and monitoring of animal diseases. *Acta Vet. Scand. Suppl.* 42, 79–85.

Olson, J. M., 1975: Experience and the Improvement of Cartographic Communication. *Cartogr. J.* 2, 97–108.

Openshaw, S., 1984: The Modifiable Areal Unit Problem Concepts and Techniques in Modern Geography. GeoBooks, Norwich, England.

Pina, M. F., S. Ferreira Alves, A. I. Correia Ribeiro, and A. Castro Olhero, 2010: Epidemiología espacial: nuevos enfoques para viejas preguntas. *Universitas Odontológica* 29, 47–65.

RKI, 2008: Infektionsepidemiologisches Jahrbuch Meldepflichtiger Krankheiten Für 2007. Robert Koch-Institut, Berlin.

RKI, 2012: Infektionsepidemiologisches Jahrbuch Meldepflichtiger KRANKHEITEN Für 2011. Robert Koch-Institut, Berlin.

Rodrigues-Silveira, R., 2013: Representación Espacial y Mapas. Centro de Investigaciones Sociológicas, Madrid.

Roth, R. E., A. W. Woodruff, and Y. F. Johnson, 2010: Value-by-alpha Maps: an alternative technique to the cartogram. *Cartogr. J.* 47, 130–140.

Rytkonen, M. J., 2004: Not all maps are equal: GIS and spatial analysis in epidemiology. *Int. J. Circumpolar Health* 63, 9–24.

Schweizer, D. M., and M. F. Goodchild, 1992: Data Quality and Choropleth Maps: An Experiment with the Use of Color. Proceedings, GIS/LIS.

Slocum, T. A., R. B. MCMaster, F. C. Kessler, and H. H. Howard, 2005: Thematic Cartography and Geographic Visualization. Pearson Prentice Hall, NJ, USA.

StatistischesBundesamt, 2014: Data on national characteristics from Statistische Ämter des Bundes und er Länder. Ergebnisse auf Grundlage des Zensus 2011. Available at: http://www.statistik-portal.de/Statistik-Portal/de_jb01_jahrtab 1.asp (accessed on 22 November 2014).

Steiniger, S., and E. Bocher, 2009: An overview on current free and open source desktop GIS developments. *Int. J. Geogr. Inf. Sci.* 23, 1345–1370.

Steiniger, S., and A. J. S. Hunter, 2013: The 2012 free and open source GIS software map – A guide to facilitate research, development, and adoption. *Comput. Environ. Urban Syst.* 39, 136–150.

Stewart, J., and P. J. Kennelly, 2010: Illuminated choropleth maps. *Ann. Assoc. Am. Geogr.* 100, 513–534.

Tversky, A., and D. Kahneman, 1974: Judgement under uncertainty: heuristics and biases. *Science* 185, 1124–1131.

Viard, T., G. Caumon, and B. Lévy, 2011: Adjacent versus coincident representations of geospatial uncertainty: which promote better decisions? *Comput. Geoscic.* 37, 511–520.

WHO, 2002: Global Surveillance of Foodborne Disease: Developing a Strategy and its Interaction with Risk Analysis Report of a WHO Consultation. World Health Organization, Geneva, Switzerland.

WHO, 2008: Foodborne Disease Outbreaks: Guidelines for Investigation and Control. World Health Organization, France.

WHO, 2014: Foodborne Disease Surveillance. Available at: http://www.who.int/foodborne_disease/en/ (accessed on 16 September 2014).

Woodruff, A. W., 2010: Value-by-alpha maps. Available at: http://andywoodruff.com/blog/value-by-alpha-maps/ (accessed on 23 October 2014).

