

# BMJ Open Funnel plots and choropleth maps in cancer risk communication: a comparison of tools for disseminating population-based incidence data to stakeholders

Walter Mazzucco,<sup>1,2</sup> Rosanna Cusimano,<sup>1,3</sup> Maurizio Zarcone,<sup>1,4</sup> Sergio Mazzola,<sup>1,4</sup> Francesco Vitale<sup>1,2,4</sup>

**To cite:** Mazzucco W, Cusimano R, Zarcone M, *et al.* Funnel plots and choropleth maps in cancer risk communication: a comparison of tools for disseminating population-based incidence data to stakeholders. *BMJ Open* 2017;**7**:e011502. doi:10.1136/bmjopen-2016-011502

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2016-011502>).

Received 3 April 2016  
Revised 28 October 2016  
Accepted 22 December 2016



CrossMark

For numbered affiliations see end of article.

## Correspondence to

Dr Walter Mazzucco;  
walter.mazzucco@unipa.it

## ABSTRACT

**Background:** Population-based cancer registries provide epidemiological cancer information, but the indicators are often too complex to be interpreted by local authorities and communities, due to numeracy and literacy limitations. The aim of this paper is to compare the commonly used visual formats to funnel plots to enable local public health authorities and communities to access valid and understandable cancer incidence data obtained at the municipal level.

**Methods:** A funnel plot representation of standardised incidence ratio (SIR) was generated for the 82 municipalities of the Palermo Province with the 2003–2011 data from the Palermo Province Cancer Registry (Sicily, Italy). The properties of the funnel plot and choropleth map methodologies were compared within the context of disseminating epidemiological data to stakeholders.

**Results:** The SIRs of all the municipalities remained within the control limits, except for Palermo city area (SIR=1.12), which was sited outside the upper control limit line of 99.8%. The Palermo Province SIRs funnel plot representation was congruent with the choropleth map generated from the same data, but the former resulted more informative as shown by the comparisons of the weaknesses and strengths of the 2 visual formats.

**Conclusions:** Funnel plot should be used as a complementary valuable tool to communicate epidemiological data of cancer registries to communities and local authorities, visually conveying an efficient and simple way to interpret cancer incidence data.

## BACKGROUND

Cancer is the second major cause of death in the developed countries.<sup>1</sup> In the past few decades, the increasing burden of disease has caused major concerns in local communities, requiring local health authorities to

## Strengths and limitations of this study

- To the best of our knowledge, this study explores for the first time the application of the funnel plot methodology to represent standardised cancer incidence ratio at the municipal level through a comparison with the commonly used visual format, as choropleth map.
- The results of this study support the use of funnel plot as a complement to choropleth map for disseminating epidemiological data of cancer registries to local communities and authorities.
- The proposed communication approach needs to be further validated in the field. To this end, the Palermo Province Cancer Registry has generated 82 municipal risk maps, one for each municipality of the province, and for a period of 1 year, qualified personnel from the registry will be involved in on-site meetings to share cancer incidence data with stakeholders (citizens, local authorities, general practitioners, specialised physicians, pharmacists, etc) using funnel plots. The Delphi consensus process will be explored as well by involving public health operators.

develop risk communication plans that address cancer incidence, survival and the potential impact of environmental exposure.<sup>2</sup> Apart from the presumed effects of lifestyle changes and environmental factors on cancer trends,<sup>3–6</sup> the global increase in cancer prevalence could be largely attributable to a combination of improved cancer survival<sup>7</sup> and ageing population.<sup>8</sup> Local communities possess a variable degree of literacy and numeracy, which, in turn, influence their understanding of such demographical and epidemiological concepts.<sup>9 10</sup> Local public health and political authorities regularly engage in finding better ways to satisfy the growing demand for information on the

impact of cancer by the general public.<sup>11</sup> In particular, citizens often question if they live in an area at high risk for environmental exposure.<sup>2</sup>

The Centers for Diseases Control and Prevention (CDC) define public health surveillance as the “Ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health.”<sup>12</sup> Population-based cancer registries (PBCRs) carry out cancer surveillance by continuously collecting and classifying information on all new cancer cases within a defined population, and providing statistics on its occurrence for the purpose of assessing and controlling the impact of this disease on the community.<sup>13</sup> The mission of PBCRs includes the translation and dissemination of evidences to enable informed decision-making and to empower the general population or other stakeholders, while preserving a rigorous methodological approach and facilitating a truthful interpretation of the data obtained. PBCR publications use validated and internationally shared measurements systems and employ terminology and visual formats that are easily understood by the scientific community, but often difficult to interpret for other stakeholders, particularly at the local level.<sup>14 15</sup>

The most commonly used format for reporting geographic comparisons of cancer epidemiological data is an atlas, which includes thematic maps, such as choropleth maps (CMs), representing cancer incidence rates (standardised rates, standardised ratios, etc) computed for specific areas.<sup>16 17</sup>

While data are available on how the context<sup>18</sup> and the content of such communications influence individual risk perception,<sup>19</sup> little is known about the effects of risk communications at a group level, particularly in small communities.<sup>20</sup>

The Italian Association of Cancer Registries (AIRtum), a national network of 41 local PBCRs, including the Palermo Province Cancer Registry (PPCR), has greatly emphasised improving communication tools.<sup>21</sup>

The aim of this paper is to propose the use of funnel plots (FPs) for reporting local cancer incidence data, as a complement to the more common visual formats employed by the PPCR to address local public health authorities and communities, in order to facilitate the dissemination and interpretation of measures of cancer statistics at the municipal level.

## METHODS

The study population consists of the 51 951 new cancer cases, excluding non-melanoma skin cancers, registered between 2003 and 2011 by the PPCR among the 1 244 239 residents of the 82 municipalities of the Palermo Province (PP; 679 850 inhabitants within the Palermo metropolitan area only).<sup>22</sup> Cancer incidence in the PP municipalities was measured by using standardised incidence ratio (SIR), defined as the ratio

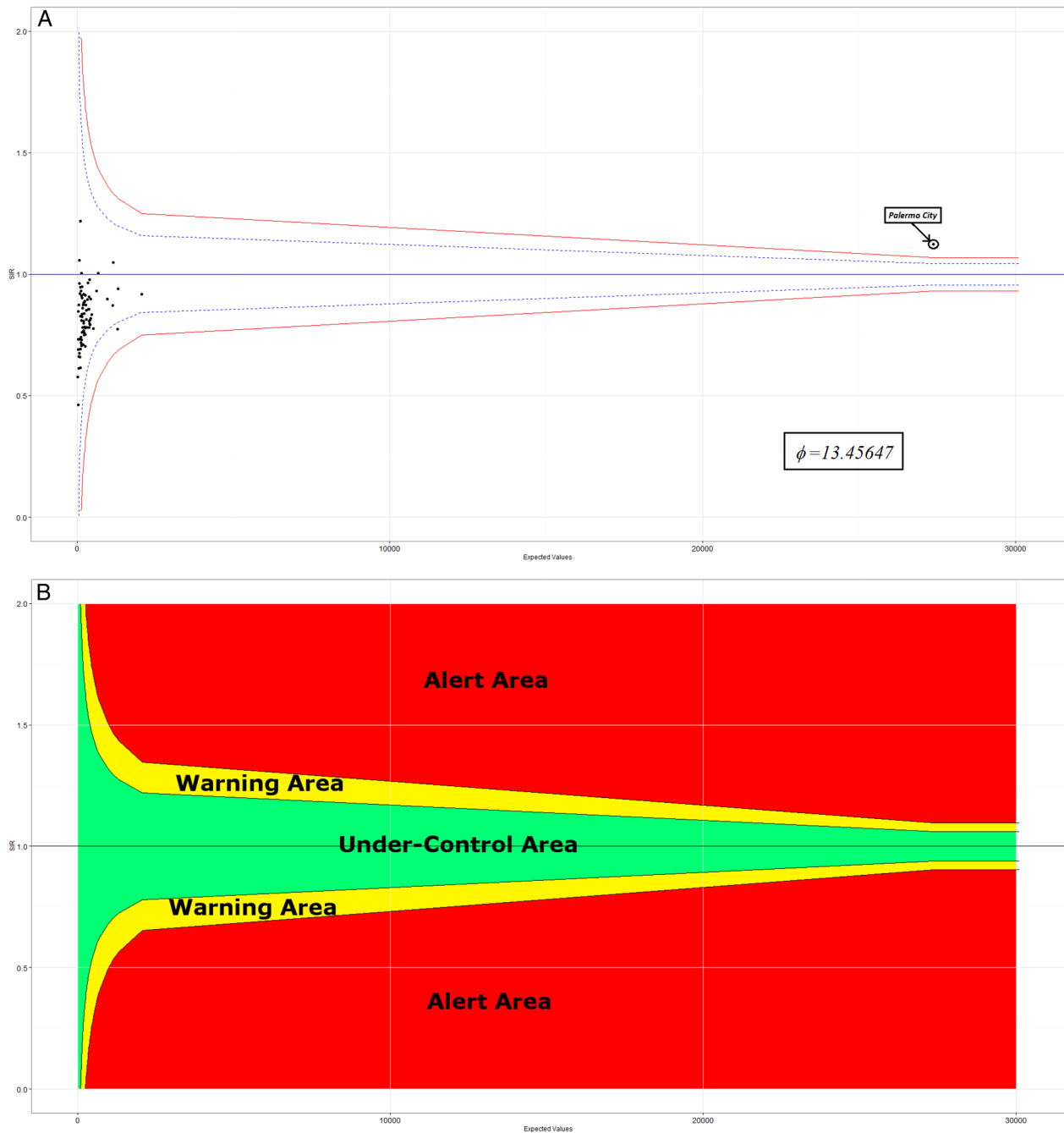
between observed cases ( $O_i$ ) and expected cases ( $E_i$ ).<sup>23</sup> The  $O_i$  were assumed to follow a homogeneous Poisson distribution with parameter  $\lambda = \theta_0 \cdot E_i$ . The  $E_i$  were estimated by indirect method,<sup>24</sup> considering the entire population time under study (the PP) as the reference population, with  $\Sigma O_i = \Sigma E_i$ .<sup>25</sup> The resident population was reported using the intercensus estimates, provided by the Italian National Statistical Institute (ISTAT), also considering the annual municipal data on migration.<sup>22</sup> For each SIR, the 95% CI was calculated by using the normal approximation method.<sup>26</sup>

Graphic FP representation<sup>26</sup> was used to highlight any municipality with a higher cancer incidence compared with the reference population (entire PP population). The following elements were included to generate the FP (figure 1A): the SIRs of the 82 municipalities, on the y-axis; the target line ( $\theta_0=1$ ), representing the reference value for the indicator of interest ( $O_i=E_i$ ); the  $E_i$  precision parameter, measuring the accuracy of the indicator of interest (Poisson variance parameter, using the hypothesis  $\theta_0=1$ ), represented on the x-axis; the 95% and 99.8% CIs, calculated with the normal approximation method, defining the control limits.<sup>26</sup> The two sets of control limit lines define three different areas within the graph (figure 1B): the ‘undercontrol’ area (in green), the ‘warning’ area (in yellow) and the ‘alert’ area (in red).<sup>27</sup>

As the data distribution was not congruent with the underlying assumption (variance equal to the expected value), in order to check for any potential overdispersion<sup>28</sup> both additive and multiplicative approaches were adopted. Overdispersion coefficients ( $\tau$  for the additive approach and  $\phi$  for the multiplicative approach) were calculated. Overdispersion was addressed by considering the winsorised estimates too.<sup>27</sup> Moreover, Z-score<sup>29</sup> and the winsorisation method (by testing for different levels of Z-score quantiles<sup>28</sup>) were applied for the direct selection of extreme values. Furthermore, to define the level of winsorisation, an R-script routine was developed to set a cut-off for the quantile between the acceptance and rejection of the overdispersion test (see online supplementary material).

The map representing the PP municipalities was generated by using the ISTAT Shapefile vector format,<sup>30</sup> released in the ED50 (European Datum - 1950) UTM Zone 32N reference system, and converted in plane coordinates (decimal degrees), providing georeferenced data in addition to the coordinates of geographic objects and their borders (for polygons), also including the information on the location of each municipality. Although traditional geographical analyses use the centroids as geostatistical units, considering that some centroid could fall outside the municipal bounds, the coordinates of the city hall were used instead.<sup>31</sup>

The PP cancer incidence variation was also shown in a CM,<sup>32</sup> representing the SIRs of each municipality. To distinguish potential high-risk and low-risk areas, a central interval of 0.95–1.05 for the colour scale was fixed, irrespective of statistical significance. Values above 1.05 and below 0.95 were divided in tertiles.<sup>33</sup>



**Figure 1** (A) Funnel plot of the SIRs in the 82 Palermo Province municipalities (study period 2003–2011); (B) cancer attention areas: ‘undercontrol’ area (in green), ‘warning’ area (in yellow) and ‘alert’ area (in red). Ninety-five per cent CIs (‘blue’ control lines) and 99.8% CIs (‘red’ control lines);  $\phi$ =overdispersion, calculated with multiplicative approach. SIR, standardised incidence ratio.

Cluster analysis was performed by using the scan statistics obtained with Openshaw’s Geographical Analysis Machine (GAM), with varying radiuses, in order to detect potentials high-risk clusters and hot spot locations, setting the p value at 0.002.<sup>34</sup> The analysis for hot spot research was performed using circles with a 3 km radius for each point of a grid, covering the study region by steps of 600 m (radius/5). The RStudio IDE (RStudio Team. RStudio: Integrated Development for R. 2015.

<http://www.rstudio.com/> (accessed 18 Jan 2016)) for the R software, V.3.1.0 (2014-04-10)—‘Spring Dance’ (R Core Team. R: A language and environment for statistical computing 2015. <http://www.R-project.org/> (accessed 18 Jan 2016)), was used to perform statistical analysis.

Finally, the weaknesses and strengths of the FP and CM methodological approaches were compared using the available literature as reference.<sup>29 33 35 36</sup>

## RESULTS

Figure 1A represents the FP of 82 municipality-specific SIRs, corrected for overdispersion ( $\phi=13.46$ ) and adjusted using the multiplicative approach.<sup>28</sup> All of the SIRs lay within the control limits, except for the Palermo city 1 (SIR=1.12), which resulted above the upper control limit line of 99.8%. Figure 1B identifies the three different cancer risk areas within the graph.

Overdispersion test results were concordant and the routine did not find out any valid value for winsorisation (see online supplementary material, section B).

Figure 2 displays the CM for cancer incidence in the 82 PP municipalities, generated by using the SIRs. The map highlights three different municipality areas (ISTAT code: 082042, 082053 and 082061; see table 1) with SIRs higher than 1.05.

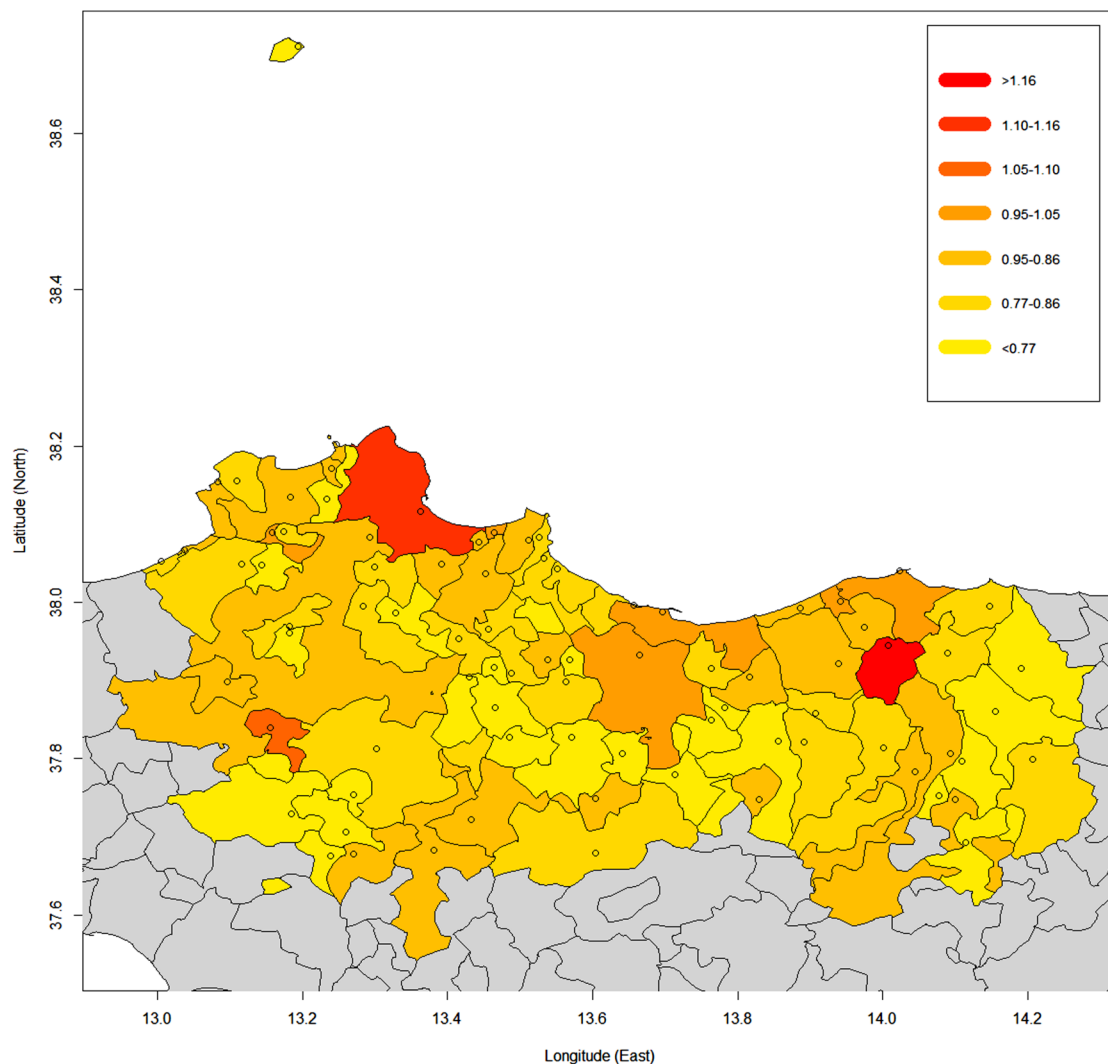
Table 1 represents the expected cases (both men and women) and SIRs with 95% CIs in the 82 PP municipalities: most of the SIRs are lower than 1 and only six municipalities present SIRs higher than 1. Among them only Palermo had a statistically significant value higher

than 1 (SIR=1, 12; 95% CIs 1.11 to 1.14) while Isnello, the municipality showing the highest SIR, failed to meet the conventional criteria for statistical significance (SIR=1.22; 95% CI 0.99 to 1.45).

No clusters were identified by the GAM approach, while a hot spot corresponding to Palermo city was highlighted (figure 3).

Table 2 summarises a comparison of the weaknesses and strengths, as per the available literature,<sup>29 33 35 36</sup> between the different visual formats explored within the context of disseminating epidemiological data to stakeholders.

As shown in the table 2, in terms of strengths, FP differed from CM in its ability to disseminate epidemiological data to stakeholders, in particular in the capability to show the scope of the phenomenon under investigation and the precision of estimates, and to highlight the significance of the estimates. On the other hand, CM, unlike FP, was able to define the spatial location of the risk and to locate the presence of any cluster. Both FP and CM were able to identify hot spots.



**Figure 2** Choropleth map of the SIRs in the 82 Palermo Province municipalities<sup>^</sup> (study period 2003–2011). <sup>^</sup>Circles represent the locations of city halls. SIR, standardised incidence ratio.

**Table 1** Expected cases and SIRs (shown in a descending order) with 95% CIs in the 82 PP municipalities (study period: 2003–2011, reference: the entire PP population)

ISTAT code	Municipality	Expected	SIR	95% CI	ISTAT code	Municipality	Expected	SIR	95% CI
082042	Isnello	104.1	1.22	0.99 to 1.45	082031	Cinisi	441.0	0.82	0.74 to 0.90
<b>082053</b>	<b>Palermo City</b>	<b>27371.4</b>	<b>1.12</b>	<b>1.11 to 1.14</b>	082007	Balestrate	292.4	0.81	0.72 to 0.91
082061	Roccamena	81.4	1.06	0.83 to 1.29	082067	Santa Flavia	414.6	0.81	0.73 to 0.89
082070	Termini Imerese	1166.4	1.05	0.99 to 1.11	082059	Pollina	148.0	0.81	0.68 to 0.94
082027	Cefalù	685.3	1.01	0.93 to 1.08	082030	Ciminna	209.5	0.81	0.70 to 0.92
082044	Lascari	152.2	1.01	0.85 to 1.17	082064	San Giuseppe Jato	379.0	0.81	0.73 to 0.89
082014	Caccamo	396.1	0.98	0.88 to 1.07	082058	Polizzi Generosa	212.6	0.80	0.69 to 0.91
082035	Ficarazzi	357.5	0.97	0.87 to 1.07	082036	Gangi	406.6	0.80	0.72 to 0.87
082038	Giardinello	84.1	0.96	0.76 to 1.17	082023	Casteldaccia	416.4	0.79	0.72 to 0.89
082056	Petralia Sottana	168.6	0.95	0.81 to 1.09	082004	Altavilla Milicia	242.8	0.78	0.68 to 0.88
082012	Bompietro	104.6	0.95	0.77 to 1.13	082005	Altofonte	379.8	0.78	0.70 to 0.86
082049	Monreale	1319.0	0.94	0.89 to 0.99	082046	Marineo	310.8	0.78	0.70 to 0.87
082052	Palazzo Adriano	123.4	0.93	0.77 to 1.10	082025	Castronovo di Sicilia	175.4	0.78	0.67 to 0.90
082079	Villabate	631.9	0.93	0.85 to 1.00	082050	Montelepre	258.1	0.78	0.68 to 0.87
082008	Baucina	102.0	0.92	0.74 to 1.10	082034	Corleone	520.3	0.78	0.71 to 0.84
082006	Bagheria	2068.7	0.92	0.88 to 0.96	082054	Partinico	1291.1	0.78	0.73 to 0.82
082076	Valledolmo	212.5	0.92	0.79 to 1.04	082051	Montemaggiore Belsito	208.5	0.77	0.66 to 0.87
082017	Campofelice di Roccella	272.6	0.91	0.80 to 1.02	082078	Vicari	156.2	0.76	0.64 to 0.88
082074	Trappeto	151.5	0.91	0.77 to 1.06	082001	Alia	225.4	0.76	0.66 to 0.86
082020	Capaci	390.4	0.92	0.82 to 0.99	082010	Bisacquino	272.7	0.75	0.66 to 0.84
082019	Camporeale	157.5	0.90	0.76 to 1.04	082063	San Cipirello	222.7	0.75	0.65 to 0.85
082071	Terrasini	446.0	0.90	0.86 to 0.98	082002	Alimena	134.9	0.74	0.62 to 0.87
082048	Misilmeri	969.4	0.90	0.84 to 0.95	082065	San Mauro Castelverde	120.1	0.74	0.61 to 0.87
082045	Lercara Friddi	335.8	0.90	0.80 to 0.99	082082	Blufi	76.5	0.73	0.57 to 0.90
082028	Cerda	243.1	0.89	0.78 to 1.01	082003	Aliminusa	73.8	0.73	0.57 to 0.90
082043	Isola delle Femmine	235.3	0.88	0.77 to 0.99	082026	Cefalà Diana	50.6	0.73	0.53 to 0.93
082032	Collesano	220.9	0.88	0.77 to 0.99	082072	Torretta	143.6	0.73	0.61 to 0.85
082024	Castellana Sicula	200.3	0.87	0.75 to 0.99	082039	Giuliana	121.2	0.72	0.59 to 0.85
082041	Gratteri	65.2	0.87	0.66 to 1.09	082047	Mezzojuso	146.5	0.71	0.60 to 0.82
082060	Prizzi	283.8	0.87	0.77 to 0.98	082055	Petralia Soprana	201.3	0.71	0.61 to 0.81
082021	Carini	1146.5	0.87	0.82 to 0.92	082062	Roccapalumba	139.0	0.71	0.59 to 0.82
082029	Chiusa Sclafani	181.4	0.86	0.73 to 0.99	082013	Borgetto	261.3	0.70	0.62 to 0.79
082009	Belmonte Mezzagno	372.3	0.86	0.77 to 0.94	082037	Geraci Siculo	112.7	0.69	0.56 to 0.82
082073	Trabia	382.2	0.86	0.77 to 0.94	082066	Santa Cristina Gela	40.6	0.69	0.48 to 0.90
082057	Piana degli Albanesi	305.0	0.85	0.76 to 0.95	082018	Campofiorito	74.2	0.67	0.52 to 0.83
082081	Scillato	36.6	0.85	0.57 to 1.12	082075	Ustica	66.6	0.66	0.50 to 0.82
082015	Caltavuturo	229.0	0.84	0.73 to 0.95	082033	Contessa Entellina	101.6	0.66	0.53 to 0.79
082080	Villafraati	166.2	0.84	0.71 to 0.96	082077	Ventimiglia di Sicilia	116.9	0.62	0.50 to 0.73
082022	Castelbuono	469.4	0.83	0.76 to 0.91	082040	Godrano	52.2	0.61	0.45 to 0.78
082068	Sciara	117.2	0.83	0.68 to 0.98	082069	Sclafani Bagni	29.4	0.58	0.37 to 0.79
082011	Bolognetta	162.6	0.82	0.70 to 0.95	082016	Campofelice di Fitalia	34.6	0.46	0.31 to 0.62

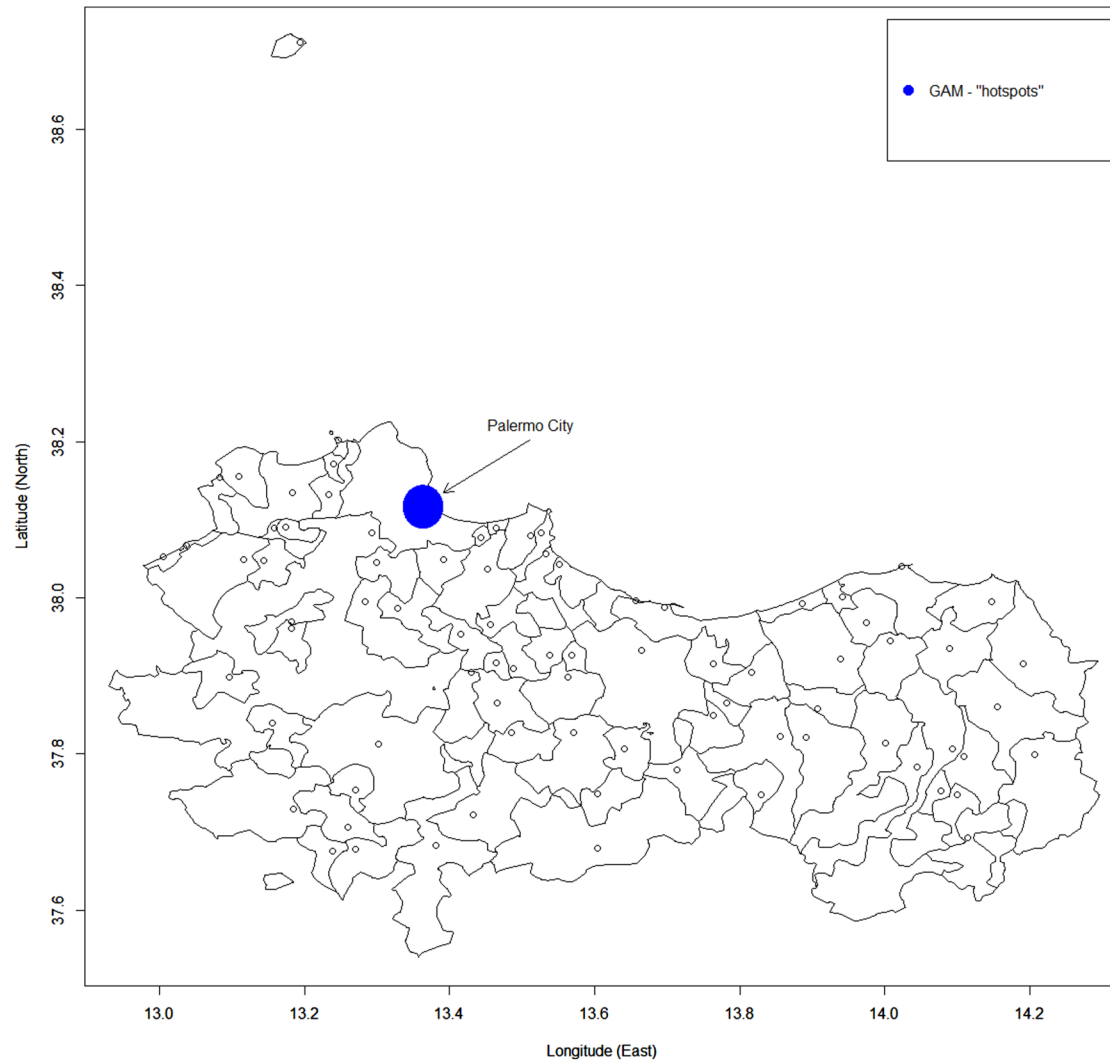
Bold typeface indicates a significant SIR value.

ISTAT, Italian National Statistical Institute; SIR, standardised incidence ratio; PP, Palermo Province.

## DISCUSSION

FPs are commonly used in process control and, in particular, in the healthcare field to compare institutional performance data;<sup>29</sup> however, this format is used for

survival<sup>37</sup> and standardised mortality ratio<sup>29</sup> in public health surveillance.<sup>38</sup> We explored the use of FPs as a supplementary tool to local provide authorities and communities with synthetic access to valid and



**Figure 3** GAM map of the Palermo Province (study period 2003–2011). GAM, Geographical Analysis Machine; SIR, standardised incidence ratio.

**Table 2** Comparison of the weaknesses (–) and strengths (+) of the funnel plot and choropleth map within the context of disseminating epidemiological data to stakeholders

Properties explored	Weaknesses and strengths of visual format	
	Funnel plot	Choropleth map
Definition of the spatial location of the risk	–	+
Identification of hot spots	+	+
Locating clusters	–	+
Displaying the scope of the phenomenon under investigation	+	–
Showing the precision of estimates	+	–
Communicating the significance of estimates	+	–

understandable cancer incidence data (SIRs) obtained at the municipal level.

Given that SIR is an effective and well-established measure in the descriptive cancer epidemiology,<sup>23</sup> we used this parameter to compare the use of FPs and the more common formats for reporting cancer epidemiological data.

Whereas scale-risk tables are easy to understand,<sup>19</sup> readers do not usually take notice of the CI, which is a

critically important measure of the precision of SIR estimates.<sup>39</sup> By displaying sample statistics together with the corresponding sample size, in relation to the control limits, FPs allow visualising both information and precision levels without the need for processing several numeric values (in this study, we used 82 point estimates and 164 confidence boundaries).<sup>38</sup> Moreover, while it is common knowledge that the numeracy skills of the general public are limited, that this obviously reduces

the general understanding of public health statistics, studies have also documented that understanding of the CIs is poor even among physicians, as heuristic reasoning often prevails on sample size.<sup>40</sup> Therefore, in order to facilitate comprehension of the epidemiological message, we have chosen the FP as a visual display method to allow the reader to identify the SIR for each municipality within the plot, and the different attention-level areas (represented by different colours) under which each location falls (figure 1B).

Reading a CM may be misleading for stakeholders<sup>41</sup> since the fear of being overexposed to environmental and other risk factors may lead to misinterpretation of the differences in colour scale, which do not properly display the potential inaccuracy in the estimation of cancer indicators (figure 2). On the other hand, the conservative choice of reporting only statistically significant increased cancer risks, as shown for the Palermo city hot spot (figure 3), excludes from the discussion the residents of most municipalities who would certainly be interested in knowing ‘what is going on in their back yard’. The combination of FP and CM, supported by tabulation of the numeric results, allows to identify locations where cancer incidence may deserve further attention, such as the municipality of Isnello, with a high SIR but a 95% CI including the null value. Clear understanding by the relevant stakeholders and their productive engagement may clarify whether such borderline findings simply reflect inadequate sample size, chance or a departure from the expected incidence that deserves further investigation.

Within the context of the chosen sample population and data, it has to be considered the presence of a single area containing a large proportion of the entire study population must be highlighted. This obviously influences each SIR value, but its potential effects are related to the study population used in the calculation of SIRs, and do not influence the FP methodology itself. Moreover, the graphic FP representation, differently from the more commonly used visual formats, allows the reader to observe, simultaneously, the situation of the municipality of interest in relation to the entire study population and to three specific areas (under control, warning and alert) representing the different attention levels. Moreover, it should also be kept in mind that the SIR values have been standardised using the EU population as external reference, allowing adjustment for age. Finally, the presence of a single area with a substantial population (Palermo city) implies an overestimation of expected cases, but the epidemiological message did not change even after the exclusion of the Palermo city area from the analysis (data not shown).

Following the methodological approach proposed, representation of the PP SIRs through FP seemed to be congruent with CM generated using the same data, with the former resulting more informative dealing with some of the dimensions explored, as shown by the comparisons of the weaknesses and strengths between the two visual formats (table 2). In particular, with regard to the

strengths of the proposed visual format, FP shows the scope of the phenomenon under investigation and the precision and significance of estimates simultaneously, by simply positioning the indicator of interest in one of the three cancer attention areas;<sup>29</sup> on the contrary, the more commonly used CMs monodimensionally represent the parameters of interest by using a different colour gradation based on the frequency distribution of the values.<sup>33 35 36</sup> The highlighted difference could be considered the main reason for making FP more comprehensive to stakeholders than CM. However, the weaknesses of FP also need to be taken into account. FP cannot be considered the ideal visual format to highlight the geographical position of the indicator of interest (SIR) and, consequently, to define any spatial cluster.<sup>29</sup> Finally, both FP and CM had the ability to identify potential hot spots, even though for CM, it is necessary to further validate the hot spot by using suitable statistical tests (eg, the GAM approach).<sup>34</sup> All of the previous considerations have led us to believe that FP could be used as a complement to CM, according to its properties, particularly in terms of validity and in terms of interpretability.

However, the proposed complementary dissemination approach needs to be further validated in the field both by involving local communities and by administering the two different visual formats to a sample of stakeholders according to the Delphi consensus process.<sup>42</sup> In fact, it can be presumed that the efficacy of a presentation format depends both on the type of format, and on the context in which the format is used (scientific vs general public).<sup>18</sup>

## CONCLUSIONS

According to the proposed comparison between the two explored methodological approaches, we concluded that FP should be considered as a complement to the current and commonly used graphical and visual formats (CMs, tables, GAM maps) to effectively communicate cancer registry statistics, particularly incidence rate, to communities and local authorities, visually conveying an efficient and simple to interpret cancer epidemiological data.

Future research on cancer risk communication should concentrate on the presentation format and on the framework in which the message is presented. From this perspective, the FP could represent a useful tool for empowering health communications to local communities and other stakeholders (patients’ associations, physicians, pharmacists, local administration, etc).

### Author affiliations

<sup>1</sup>Registro Tumori di Palermo e Provincia, University of Palermo, Palermo, Italy

<sup>2</sup>Dipartimento Scienze Promozione della Salute e Materno Infantile

“G. D’Alessandro”, Sezione di Igiene, University of Palermo, Palermo, Italy

<sup>3</sup>Azienda Sanitaria Provinciale 6 di Palermo—Dipartimento di Prevenzione, Palermo, Italy

<sup>4</sup>Unità Operativa Complessa di Epidemiologia Clinica con Registro Tumori, Azienda Ospedaliera Universitaria Policlinico “Paolo Giaccone”, Palermo, Italy

**Contributors** All individuals listed as authors have contributed substantially to designing, performing or reporting of the study and every specific contribution is indicated as follows. WM, RC, MZ and SM were involved in conception and design of the study. MZ and SM were involved in statistical analysis. WM, RC, MZ and SM were involved in interpretation of data. WM and RC were involved in manuscript writing and drafting. FV, WM and RC were involved in revision of the manuscript. WM, RC, MZ, SM and FV were involved in approval of the final version of the manuscript. The document has been reviewed and corrected by a native English speaker with extensive scientific editorial experience to ensure a high level of spelling, grammar and punctuation.

**Funding** This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** Online supplementary data (results of overdispersion tests, R-script to detect the greatest cut-off for the winsorisation procedure) have been provided as an online supplementary file. Other statistical results are available by emailing [walter.mazzucco@unipa.it](mailto:walter.mazzucco@unipa.it).

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

- Summary tables of mortality estimates by cause, age and sex, globally and by region, 2000–2015. In: Estimates for 2000–2015 - Health statistics and information systems WHO. [http://www.who.int/entity/healthinfo/global\\_burden\\_disease/GHE2015\\_Deaths\\_Global\\_2000\\_2015.xls?ua=1](http://www.who.int/entity/healthinfo/global_burden_disease/GHE2015_Deaths_Global_2000_2015.xls?ua=1) (accessed 22 Mar 2017).
- Barchitta M, Fragapane S, Quattrocchi A, et al. Environmental health risk communication in the case “Terra dei Fuochi”: content analysis of online newspaper articles. *Ann Ig* 2015;27:30–8.
- Crocetti E, Pirastu R, Buzzoni C, et al. Results in 3 Rapporto Sentieri. Studio epidemiologico nazionale dei territori e degli insediamenti esposti a rischio da inquinamento. *Epidemiol Prev* 2014;38(Suppl 1): 29–124.
- Straif K, Cohen A, Samet J. *Air pollution and cancer*. IARC Scientific publication no.161. <https://www.iarc.fr/en/publications/books/sp161/AirPollutionandCancer161.pdf> (accessed 18 Jan 2016).
- Agents classified by the IARC monographs, volumes 1–117. In: Monographs on the evaluation of carcinogenic risk to humans. International Agency for Research on Cancer (IARC). <http://monographs.iarc.fr/ENG/Classification/> (accessed 22 Mar 2017).
- Boffetta P, Nyberg F. Contribution of environmental factors to cancer risk. *Br Med Bull* 2003;68:71–94.
- Allemani C, Weir HK, Carreira H, et al. Global surveillance of cancer survival 1995–2009: analysis of individual data for 25,676,887 patients from 279 population-based registries in 67 countries (CONCORD-2). *Lancet* 2015;385:977–1010.
- Alwan A, Armstrong T, Betcher D, et al. WHO Global Status Report on non-communicable diseases 2010. Geneva, 2011.
- Melissa G. *Health literacy and numeracy. Workshop summary*. Washington: The National Academy Press, 2014. <http://www.nap.edu/catalog/18660/health-literacy-and-numeracy-workshop-summary> (accessed 18 Jan 2016).
- Reyna VF, Nelson WL, Han PK, et al. How numeracy influences risk comprehension and medical decision making. *Psychol Bull* 2009;135:943–73.
- Edwards A, Elwyn G, Mulley A. Explaining risks: turning numerical data into meaningful pictures. *BMJ* 2002;324:827–30.
- CDC. Updated Guidelines for evaluating Public Health Surveillance Systems. *MMWR* 2001;50(RR13):1–35. <http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm> (accessed 18 Jan 2016).
- Bray F, Soerjomataram I. The changing global burden of cancer: transitions in human development and implications for cancer prevention and control. In: Gelband H, Jha P, Sankaranarayanan R, Horton S, eds. *Cancer: disease control priorities*. 3rd edn. Vol 3. Washington DC: The International Bank for Reconstruction and Development/The World Bank, 2015. Chapter 2, pp 23–44.
- Rapporto sui tumori a Palermo e provincia. Anni 2003–2007. Registro Tumori Palermo e Provincia. Anno 2011. <http://www.unipa.it/dipartimenti/dipartimentodiscienzeperlapromozionedellasaluteematernoinfantileg.dalessandro/content/documenti/Rapporto-sui-tumori-a-Palermo-e-Provincia-anni-2003-2007-1.pdf> (accessed 18 Jan 2016).
- Cancer in Friuli Venezia Giulia Incidence, Survival and Prevalence: 1995–2005. Agenzia Regionale della Sanità Friuli Venezia Giulia. [http://www.cro.it/opencms/export/sites/cro/it/azienda\\_informa/\\_docs/registro-tumori-1995-2005.pdf](http://www.cro.it/opencms/export/sites/cro/it/azienda_informa/_docs/registro-tumori-1995-2005.pdf) (accessed 18 Jan 2016).
- Jemal A, Vineis P, Bray F, et al. *The Cancer Atlas*. 2nd edn. Lyon: Non series publications IARC, 2014.
- Atlante Sanitario della Sicilia. Suppl. monografico Notiziario Dipartimento per le Attività Sanitarie ed Osservatorio Epidemiologico. Assessorato della Salute, Regione Sicilia. Anno 2013. [http://pti.regione.sicilia.it/portal/page/portal/PIR\\_PORTALE/PIR\\_LaStrutturaRegionale/PIR\\_AssessoratoSalute/PIR\\_AreeTematiche/PIR\\_Atlanteoncologico](http://pti.regione.sicilia.it/portal/page/portal/PIR_PORTALE/PIR_LaStrutturaRegionale/PIR_AssessoratoSalute/PIR_AreeTematiche/PIR_Atlanteoncologico) (accessed 18 Jan 2016).
- Visschers VH, Meertens RM, Passchier WW, et al. Probability information in risk communication: a review of the research literature. *Risk Anal* 2009;29:267–87.
- Lipkus IM, Hollands JG. The visual communication of risk. *Natl Cancer Inst Monogr* 1999;25:149–63.
- McInerney GJ, Chen M, Freeman R, et al. Information visualization for science and policy: engaging users and avoiding bias. *Trends Ecol Evol* 2014;29:148–57.
- Mappa dei Registri Tumori di popolazione. Associazione Italiana Registri Tumori (AIRTUM) 2015. <http://www.registri-tumori.it/cms/it/copertura> (accessed 18 Jan 2016).
- Ricostruzione intercensuaria della popolazione 2002–2011. Istituto Nazionale di Statistica. 2015. <http://www.demo.istat.it> (accessed 18 Jan 2016).
- Greenland S, Rothman KJ, Lash TL. Measures of effect and measures of association. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern epidemiology*. 3rd edn. Philadelphia: Lippincott Williams & Wilkins, 2008:67–9.
- Mantel N, Stark CR. Computation of indirect-adjusted rates in the presence of confounding. *Biometrics* 1968;24:997–1005.
- Morris SE, Wakefield JC. Assessment of disease risk in relation to a pre-specified source. In: Elliot P, Wakefield JC, Best NG, Briggs DJ, eds. *Spatial epidemiology: methods and applications*. New York: Oxford University Press, 2000:53–184.
- Spiegelhalter DJ. Funnel plots for institutional comparisons. *Qual Saf Health Care* 2002;11:390–1.
- Spiegelhalter DJ. Handling over-dispersion of performance indicators. *Qual Saf Health Care* 2005;14:347–51.
- Silcocks P. Hazard ratio funnel plots for Survival. *J Epidemiol Community Health* 2009;63:856–61.
- Spiegelhalter DJ. Funnel plots for comparing institutional performance. *Stat Med* 2005;24:1185–202.
- Confini amministrativi, dei sistemi locali del lavoro e delle NUTS2 in versione generalizzata. Istituto Nazionale di Statistica 2010. [http://www.istat.it/it/files/2011/04/com2011\\_g.zip](http://www.istat.it/it/files/2011/04/com2011_g.zip) (accessed 18 Jan 2016).
- Mazumdar S. Shape and scale in detecting disease clusters. [PhD (Doctor of Philosophy) thesis]. University of Iowa, 2008. <http://ir.uiowa.edu/etd/208> (accessed 18 Jan 2016).
- Richards TB, Berkowitz Z, Thomas CC, et al. Choropleth map design for cancer incidence, part 2. *Prev Chronic Dis* 2010;7:A24.
- Cromley G, Cromley EK. Choropleth map legend design for visualizing community health disparities. *Int J Health Geogr* 2009;8:52.
- Besag J, Newell J. The detection of clusters in rare diseases. *Wiley R Stat Soc* 1991;154:143–55.
- Sui DZ, Holt JB. Visualizing and analysing public-health data using value-by-area cartograms: toward a new synthetic framework. *Cartographica* 2008;43:3–20.
- Kumar N. Frequency histogram legend in the choropleth map: a substitute to traditional legends. *Cartogr Geogr Inform Sci* 2004;31:217–36.
- Quaresma M, Coleman MP, Rachet B. Funnel plots for population-based cancer survival: principles, methods and applications. *Stat Med* 2014;33:1070–80.
- Dover DC, Schopflocher DP. Using funnel plot in public health surveillance. *Popul Health Metr* 2011;9:58.
- Hildon Z, Allwood D, Black N. Impact of format and content of visual display of data on comprehension, choice and preference: a systematic review. *Int J Qual Health Care* 2012;24:55–64.
- Rakow T, Wright RJ, Spiegelhalter D, et al. The pros and cons of funnel plots as an aid to risk communication and patient decision making. *Br J Psychol* 2015;106:327–48.
- Ocana-Riola R. Common errors in disease mapping. *Geospat Health* 2010;4:139–54.
- Hsu CC, Brian A. The Delphi technique: making sense of consensus. *Pract Assess Res Eval*. *Pract Assess Res Eval* 2007;12:1–8.