A data narrative about tuberculosis pandemic in Gabon

Raymond Ondzigue Mbenga[1,2], Veronika Peralta[1], Thomas Devogele[1], Faten El Outa[1], Sydney Maghendji Nzondo[2] and Edgard Brice Ngoungou[2]

[1]LIFAT, University of Tours, Blois, France
[2]DEBIM-UREMCSE, Univ. of Health Sciences, Libreville, Gabon

Abstract
Data narration is the activity of crafting narratives supported by facts extracted from data analysis, using interactive visualizations. It allows the transmission of findings in the data, by visual means, in order to facilitate their reception by a target audience. Despite its recognized utility in public health, data narratives are typically limited to the transmission of treatment recommendations to educate the general public. This paper describes the crafting of a data narrative about tuberculosis pandemic in Gabon, intended to an audience of health professionals and authorities. Specifically, we describe and illustrate all phases of the crafting process, combining best practices in data narration and epidemic intelligence.

Keywords
Data narrative, tuberculosis, Gabon

1. Introduction

Data narration, i.e. narrating with data visualisations [1], is receiving increasing interest in several communities such as journalism, business, e-government and data science. A data narrative is defined as a structured composition of messages, which convey findings over the data, and are typically delivered via visual means in order to facilitate their reception by a intended audience [2]. More specifically, data narratives can be seen as ordered sequences of steps, each of which may contain words, images, visualisations, audio, video, or any combination thereof, and which are based on data [3].

Despite its recognised usefulness in public health [4], where information on health problems must compete with thousands of other communication messages (infodemics), data narratives are little used, and are typically limited to the transmission of treatment recommendations and patient feedback to educate the general public. Few works aim at conveying scientific results to an expert audience to report on the health situation and the results of implemented policies, and more generally, to help decision making. Furthermore, such data narratives are build following ad-hoc processes. In a general way, there is a lack of processes and methodological guidelines for data narration in public health.

In this paper we describe the crafting of a data narrative about tuberculosis (TB) pandemic in Gabon. This narrative is intended for health authorities and experts in Epidemic Intelligence (EI), with the goal of describing the epidemiological situation of tuberculosis in a pilot area of study, the Libreville-Owendo-Akanda (LBVO-A) health region, before a countrywide move.

Specifically, we describe, for each step of the process, the key methodological aspects, the major difficulties and the particularities of epidemiological domain. Indeed, the crafting process customizes general data narration ones incorporating specific features of epidemiology, best practices in EI and communication to epidemiologists and health authorities. In particular, data collection, processing and analysis have a key place in the process, such tasks being at the core of EI. In addition, classical statistic analysis is enriched with other data mining tasks and confronted to the state of the art, the latter being a specific requirement when addressing to a scientific audience. From a technical perspective, the underlying system supports spatio-temporal data of very heterogeneous quality. Indeed, geographic information systems are well-suited for EI applications [5].

The main applied and methodological contributions are:

- A detailed description of the crafting process, from the collection of analytical needs to the visual rendering of results.
- A discussion of the main challenges in data narration in the field of epidemiology.
- A summary of the main messages learned from the data, which form the heart of the narrative.
- An example of usage of data narration as a decision support tool to enable decision makers to define a better tuberculosis control strategy.

Published in the Workshop Proceedings of the EDBT/ICDT 2022 Joint Conference (March 29-April 1, 2022), Edinburgh, UK
raymond.ondziguembenga@etu.univ-tours.fr (R. Ondzigue Mbenga); veronika.peralta@univ-tours.fr (V. Peralta); thomas.devogele@univ-tours.fr (T. Devogele); faten.elouta@etu.univ-tours.fr (F. E. Outa); sydneymaghendji@yahoo.fr (S. Maghendji Nzondo); ngoungou2001@yahoo.fr (E. B. Ngoungou)
0000-0002-0877-7063 (R. Ondzigue Mbenga); 0000-0001-7116-9338 (V. Peralta); 0000-0002-9421-8566 (T. Devogele)

The paper is organized as follows. Section 2 describes related work on data narration and epidemic intelligence processes. Section 3 describes the main phases of the crafting process and the obtained results. Finally, Section 4 presents lessons learned and Section 5 concludes.

2. Related work

In this sub-section we provide some background on tuberculosis and we describe related works on data narration and EI processes. We also review the use of data narration in public health.

2.1. About tuberculosis

Tuberculosis is an infectious disease caused by a bacterium, Mycobacterium Tuberculosis, which is transmitted by air when an infected person coughs or sneezes. In the most frequent form of TB (pulmonary TB) the pathogen attacks the lungs, but in more complex forms (extra-pulmonary TB) it can also affect other organs (e.g.lymph nodes, bones, kidneys, etc.). Typical symptoms of TB are fever with night sweats, chronic cough, fatigue, shortness of breath and loss of appetite.

Tuberculosis continues to be a serious public health problem whose magnitude requires global attention. According to the World Health Organization (WHO), this pandemic infected 10 million people around the world in 2019. In the African region, that of Sub-Saharan Africa is the most infected with nearly two million cases reported each year. Among these countries, Gabon is the second most affected with an annual incidence of 428 cases per 100,000 inhabitants, behind Zimbabwe (562 cases per 100,000 inhabitants). The LBVO-A health region alone accounts for more than 50% of tuberculosis cases notified in Gabon.

2.2. Data narration process

Data narration, i.e. the crafting of a data narrative, is a complex process at the crossroads of several domains : data processing, data analysis, data visualisation, communication, among others. Despite numerous contributions in each of these fields, few works propose global methodologies, describing the whole data narration process.

Firstly, Chen et al. [6] distinguished two main phases typically used for data narration : (a) *visual analytics*, which requires to see all aspects of complex data, explore their interrelationships, and is supported by multiple coordinated views and sophisticated interaction techniques, from (b) *storytelling*, which is meant to convey only interesting and/or important information extracted through the analysis, presented in a simple and easily understandable way. They proposed an intermediate phase, *data synthesis*, in which the analyst assembles and organizes information pieces to be communicated, facilitating the telling of visual analytical findings in a compelling narrative.

The same three phases (but named differently) are described by Lee et al. [7]: (i) *explore data*, to retrieve findings among data, (ii) *make a story*, to turn findings into a sequence of narrative pieces and build the plot of the narrative, and (iii) *tell a story*, to render the plot via visual means.

Recently, a conceptual model of data narrative [2] formalises it main concepts and their relationships. The model aims to guide the author through the data narration process: define the analysis *goals* and break them down into *analytical questions*, collect and explore data through a set of *collectors* that allow to manipulate data with varied tools for answering analytical questions, analyze data and underline the *findings*, bring out *messages* (from the findings) to communicate to the audience, structure messages to build the narrative in terms of *acts* and *episodes* and render it with visual means via *dashboards*. Even if the conceptual model can be instantiated by different processes (with different workflows), in a demonstration scenario [8], the authors illustrate a particular proceeding in 4 phases : (i) *goal setting*, (ii) *data exploration*, (iii) *narrative organisation*, and (iv) *presentation*.

We highlight that despite naming discrepancies, the described proposals agree in the major phases of the data narration process.

2.3. Epidemic intelligence process

The World Health Organization (WHO) proposed an EI process [9], with several protocols depending on whether data is collected through Indicator-based surveillance (IBS) or Event-based surveillance (EBS). While the former deals with data that has been previously validated, the latter focuses on new data, mostly based on rumours and unverified information, therefore requiring a validation process. Given the retrospective character of data narration and the importance of data validation, we consider only IBS protocols, relaying on sources in the health sector.

The process is organized into five main phases: (i) *detection of raw data*, which consists in selecting data sources and collecting data, (ii) *triage of relevant data and information*, which concerns data analysis (data quality and descriptive and analytical epidemiology) and data interpretation (qualitative assessment of the significance of findings), (iii) *verification of signal*, which consists in confirming the authenticity and conformity of the findings and their characteristics, generally by cross-checking using other reliable sources, (iv) *risk assessment of the event*, which implies determining the level of risk to human health and the potential control measures that can be implemented, and (v) *communication*, which concerns the communication of indicators to different audiences (experts, health authorities, partners, population, etc.) to help in the decision-making process.

Apart from the communication phase, the rest of the phases of the EI process proposed by WHO correspond to the data exploration phase of the data narration process. Indeed, nothing is advised concerning the way of crafting messages (explaining the learned and validated signals) and structuring them in a coherent discourse. Finally, the communication phase of the EI process corresponds to the presentation phase of the data narration process, but also to effective dissemination of the information. In particular, the WHO suggests that various supports may be used to share information, but the use of visual artifacts is not detailed.

2.4. Data narration in public health

Many works describe the use of data narration techniques in public health. They mainly concern the communication phase, dealing with storytelling issues.

In the field of public health, there are many communication strategies, but the most effective is storytelling [10].

There is a growing trend to use storytelling as a research and intervention tool on public health issues, especially those with a strong disease prevention component [11].

Indeed, storytelling is used to educate populations on health protection practices, to advocate for improved clinical care and to encourage efforts to combat infectious diseases [12].

For example, in [13], authors tested a new communication modality to promote prevention messages and colorectal cancer screening among Latinos. In [14], authors conducted a survey to determine the effectiveness of a rational emotional digital storytelling therapy on HIV/AIDS knowledge and risk perception among school children in Enugu State, Nigeria.

In [15], authors developed a digital storytelling intervention on diabetes to raise awareness among immigrant and refugee populations with limited English proficiency.

In another study conducted in rural Alaskan communities [16], authors investigated whether digital stories could influence participants' feelings about cancer, and whether viewing the digital stories led to a change (or intention to change) in health behavior.

These studies showed that storytelling can increase patients' positive responses and increase their level of knowledge.

To address the need for compelling and successful information visualizations in biomedical sciences, authors of [17] propose a theoretical framework for visual storytelling and illustrate its potential application through two visual stories, one on vaccine safety and the other on cancer immunotherapy. Both examples, based on data, combine multiple media (photographs, illustrations, choropleth maps, tables, graphs, and diagrams) with text to create powerful visual stories for selected target audiences.

A COVID-19 monitoring dashboard that provides data exploration and visualization capabilities was designed in [18]. One of its applications is to support data storytelling.

According to the state of the art, data storytelling techniques have been used many times in the field of public health, both applied to population awareness and education, and recently also to EI. The proposed data narratives were crafted in an ad-hoc way, not following neither EI nor data narration processes, and only the presentation phase (data storytelling) was reported. To the best of our knowledge, this article is the first to describe the complete process of crafting a data narrative about public health phenomena.

3. Crafting of a data narrative about tuberculosis

In this section we describe the crafting of a data narrative to tell tuberculosis pandemic situation to health authorities in Gabon. The crafting process is inspired by classical data narration ones [6, 7, 2] (described in Subsection 2.2), and enriched with particular tasks from the WHO EI process [9]. The following sub-sections describe each of the phases and the obtained results.

3.1. Goal setting

In this phase, the goal of the data narrative is defined. From this objective, one or more analytical questions can be derived.

Objective definition The objective of this data narrative is to describe the epidemiological situation of tuberculosis in the LBVO-A health region of Gabon between 2016 and 2018. Specifically, it aims to provide answers to decision-makers on the epidemiological profile of the disease, in order to enable them to better orient their decisions for an effective response.

Analytical questions formulation By refining this objective and conducting interviews with various officials, we obtained the following initial list of analytical questions.

- Q1: What are the epidemiological characteristics of tuberculosis? We aim to describe the profile (frequency, variations) of tuberculosis according the characteristics of the tuberculosis patients

Q2: What is the spatial and temporal distribution patients? Identifying the most affected areas essential to target response actions.

Other analytical questions may arise during data ploration.

3.2. Data exploration

Data exploration consists in exploring an epidemiological dataset for trends, patterns and correlations. Considering tasks proposed in EI and data narration processes, we proceed in 7 steps: (i) data collection, (ii) data processing (to solve data quality problems), (iii) data analysis (descriptive and analytical epidemiology), (iv) data interpretation (qualitative assessment of findings),

(v) verification (cross-checking and comparison with the state of the art), (vi) epidemic risk assessment, and (vii) message formulation (key messages, based on findings, to the target audience). This steps are no supposed to be executed sequentially, and many back and forth transitions may be necessary.

Data collection Sociodemographic and clinical data were collected from 7968 medical records of patients undergoing TB treatment at the Nkembo Specialized Hospital in Libreville. They were completed with geographical data about the administrative boundaries (neighbourhoods and districts) of the LBVO-A region.

Patients and their treatments are described in terms of the following dimensions:

- *Time*: The year of the treatment (2016 to 2018);
- *Type of TB*: Clinical form (pulmonary, extra-pulmonary, multidrug-resistant or unknown);
- *Age*: The age of the patient, organised in two levels: age and age group;
- *Gender*: The gender of the patient (male or female);
- *Profession*: The profession status of the patient;
- *Geography*: The patient's place of residence, organised in two levels: neighbourhood and district;
- *HIV status*: The patient's HIV status, whether tested (negative, positive) or unknown;
- *Treatment outcome* (cured, completed, failure, deceased, lost to follow-up[1], discontinued and transferred).

[1] According to the WHO, a "lost to follow-up" is a patient whose treatment has been interrupted for at least two consecutive months and for whom no treatment outcome has been assigned (including patients transferred to another treatment unit and those whose treatment outcome is not known)

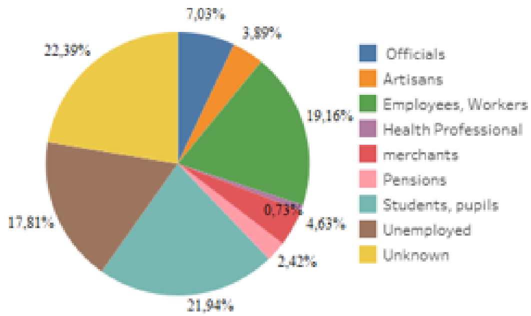

Figure 1: Distribution of cases by professional status

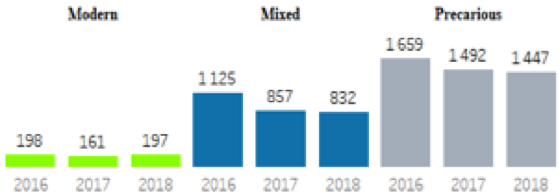

Figure 2: Evolution of cases by type of neighbourhood

Data processing Collected data were processed in order to solve several data quality problems, including disambiguation of empty fields, standardization of data types, correction of geographic boundaries (e.g. for overlapping districts). Age groups were computed and professional status were aggregated.

Clean data were stored in a PostgreSQL spatial data warehouse.

Data analysis Data were analyzed using the interactive BI tool Tableau Desktop[2], devising many queries and noting findings.

Below, we present four examples of the analyzes carried out. Since tuberculosis mostly affects vulnerable people, we began by studying univariate distributions of the dimensions most concerned by this aspect (profession, age, HIV status and geography). Findings set off new queries.

The distribution of cases by professional status, illustrated in Figure 1, shows a predominance of students (21.94%) and unemployed patients (17.81%). The proportion of patients with unknown profession is also high (22.39%).

The distribution of cases by age group (see top part of Figure 3) shows more patients among young adults (20-34 years) and mature adults (35-64 years).

The spatial and spatio-temporal distributions of cases by district and by neighbourhood (figures omitted because of their size) did not show any spatial correlation,

[2] https://www.tableau.com/

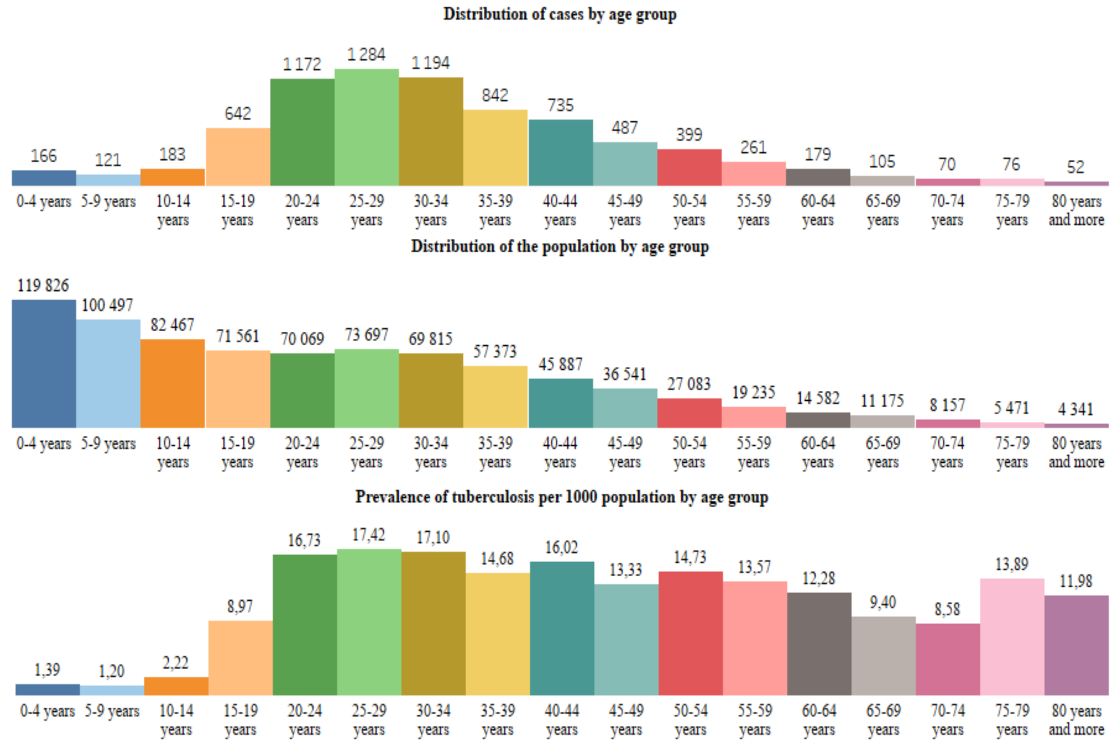

Figure 3: Distribution of cases by age group (top); age pyramid in the study region (middle); prevalence by age group, over 3 years, per 1000 inhabitants (bottom)

which is very common in other spatial epidemiology studies. In order to investigate further, we studied the distribution by neighbourhood typology (see Figure 2), where the trends are more marked (p-value = 0.01). Additional data sources were selected for this analysis, and additional statistical tools were used.

The distribution of cases by treatment outcome revealed that 74.60% of patients are lost to follow-up. This alarming finding lead to a new analytical question:

Q3: What are the epidemiological characteristics of patients lost to follow-up?

Data interpretation Hereafter, we present a detailed example of data interpretation.

The high proportion of patients being students could be linked to the high consumption of drugs in schools. Indeed, a study carried out at the Paul Idjendjé Gondjout School in Libreville [19] showed that among the surveyed pupils, aged between 14 and 20 years, 40% consumed alcohol, 20% smoked tobacco and 15% smoked cannabis.

To put these findings into perspective, given relatively young population of Gabon, we studied the age pyramid in the study area (middle part of Figure 3), taken from the Gabon's 2013 General Population and Housing Census (RGPL 2013), which evidences that more than half of the population is under 22 years old.

These two results allow the calculation of prevalence (number of patients in the study period per 1000 inhabitants) by age group (bottom part of Figure 3). It can be seen that children are little affected by tuberculosis, but young adults are only slightly more affected than older adults. Finally, it should be noted that the small number of senior patients does not allow the formulation of statistically-valid insights.

Verification Analysis results showed that more than half of the patients (57.71%) come from poor neighbourhoods (see Figure 2). This result is in line with those of [20].

Other results (not described in data analysis paragraphs above) were also verified. For example, men are more affected by tuberculosis (61.60% of patients), which is confirmed by other studies [21, 22]; pulmonary forms of TB were observed in a large majority of patients (90%) and drug-resistant forms were marginal (0.28%), which is in line with professional knowledge.

Other results deviate from the state of the art. For example, the proportion of patients lost to follow-up (75%) is too high compared to other countries in the Sub-Saharan African sub-region (e.g. in Mali [23]). This discrepancy must be also informed. In addition, the proportion of unemployed patients (17.81%), even being substantial, is very low compared to those found in other studies [24, 25]. This proportion must be completed by the high number of patients with unknown profession (see Figure 1).

Epidemic risk assessment There is a risk of spreading tuberculosis among the student population.

Patients lost to follow-up may lead to a generalized TB epidemic in the country. Geographically disadvantaged areas are more risky.

Message formulation The findings obtained during data analysis and interpretation, after validation and risk analysis, allow the formulation of the following messages (exhaustive list) for the audience:

- M1: Apart from patients with unknown profession (22.39%), students are the most affected by TB (21.94%). The influence of drug use is a lead to be explored.
- M2: The proportion of unemployed patients (17.81%) is lower than in other African countries.
- M3: A high proportion of patients (45.81%) are between 20 and 34 years old.
- M4: The highest prevalence values (from 16.73 to 17.42 per 1000 inhabitants, in 3 years) also happens for patients between 20 and 34 years old, but prevalence slightly drops for older patients.
- M5: Children are little affected by TB (prevalence is lower than 2.22 per 1000 inhabitants, in 3 years).
- M6: The proportion of patients lost to follow-up (74.60%) is alarming. It is higher than in other African countries.
- M7: There is no spatial nor spatio-temporal correlation.
- M8: The spatial and temporal evolution is correlated with the typology of the neighbourhoods (p-value = 0.01), with precarious and mixed neighbourhoods being significantly more affected, but showing a slight downward trend.
- M9: A very large majority of patients (90%) suffer from the pulmonary form, which is consistent with professional knowledge. Multidrug-resistant TB has been recorded in 0.28% of patients.
- M10: The distribution by clinical form of those lost to follow-up is very similar to that of all patients.
- M11: Among the patients, there is a male predominance (61.60%).
- M12: The proportion of patients with unknown HIV status is very high (65%). This does not allow this criterion to be considered in the patient profile.
- M13: The sixth district of Libreville records the highest proportion of cases (27.64%).
- M14: In all the districts, the general trend is a slight drop in prevalence. However, two districts (1st of Akanda and 2nd of Owendo) show particular trends and greater variations, partialy explained by their small populations.
- M15: There are two hot districts for TB, Les-PK and Nzeng-Ayong, with respectively 13.% and 11.58% of patients.
- M16: More than half of the patients (57.71%) come from poor neighbourhoods.
- M17: In the distribution of patients lost to follow-up by type of neighbourhoods, we find the same trend as for all patients.
- M18: TB affects all social classes and genders, with most patients being adults and living in poor neighbourhoods.
- M19: The profile of the patients lost ot follow-up, in terms of gender, age, occupation and location, is very similar to that of other patients.

3.3. Narrative structuring

In this phase, messages are assembled and ordered in a coherent and logical plot in order to facilitate their understanding and attract the audience. We started by setting the audience and selecting the messages to convey to such audience. Then, we choose the narrative structure and organized messages.

Audience choice The target audience for this TB data narrative is composed of health authorities in Gabon, including the head of the National Tuberculosis Control Program (PNLT), the head of the Institute of Epidemiology and Endemic Control (IELE), policy makers, as well as epidemiologists and doctors involved in the fight against TB.

Message selection All messages formulated during data exploration are selected to be communicated to the target audience.

Structure choice Following data narration recommendations [2], we organize the plot of the narrative in several acts. Acts are composed of episodes, each one telling a message.

The plot is organized in 8 acts, listed in the first columns of Table 1.

The first act introduces the study intention. The second act presents the salient messages and the next 6 acts focus on a dimension describing the patients, respectively, occupation, gender, age, HIV status and geography. The final act presents conclusions and recommendations.

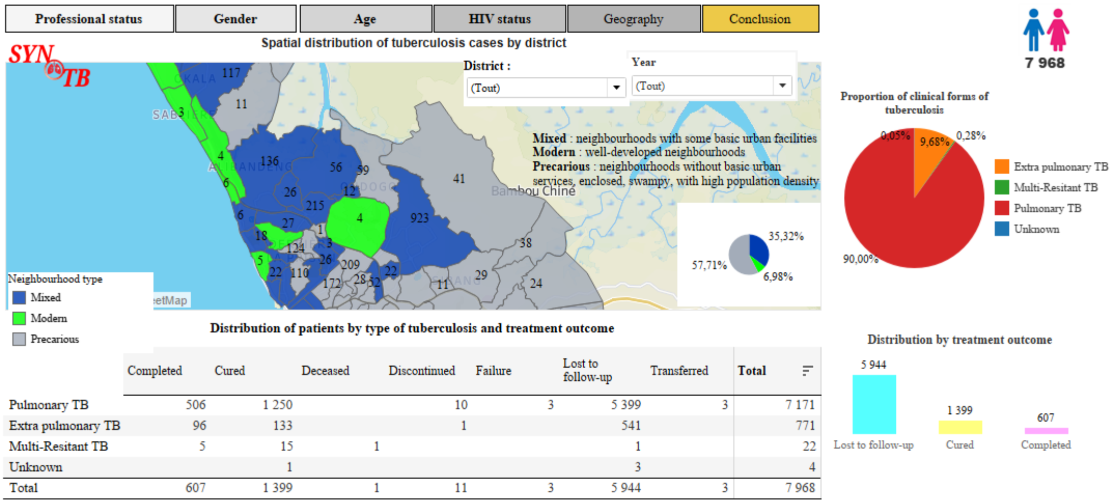

Figure 4: Dashboard for Act II

Act	Title	Included messages
I	Introduction	
II	General presentation	M6, M9, M10, M16
III	Profession	M1, M2
IV	Gender	M11
V	Age	M3, M4, M5
VI	HIV status	M12
VII	Geography	M7,M8,M13,M14,M15, M17
VIII	Conclusion	M18, M19

Table 1
Structuring the narrative

We have opted for this structuring, as it facilitates the understanding of a patient profile, according to the various dimensions that make it up. It is intended that the audience should first look at the introduction and general presentation, but then they should be able to navigate between the following acts according to their needs. The order of navigation has no impact on the conclusions.

This structure is well known in modern storytelling as Martini-Glass [26]. It combines several types of interactivity in a balanced way: The author lays down his narrative path (Acts I and II) and the reader interacts and explores the available paths to better understand the data (Acts III to VIII). This makes navigation through the story flexible, according to the needs of the audience.

Message mapping Finally, messages were mapped to acts, as shown in Table 1.

3.4. Presentation

This phase concerns the choice of the visual representation (e.g. interactive dashboards, infographics, slideshow, video) and the setting of visual artifacts (graphics, colors, text, etc.) for telling acts and episodes, attracting the audience attention.

Visual representation choice We designed and implemented two versions of the data narrative with different visual rendering: (i) an interactive narrative, composed of interconnected interactive dashboards, and (ii) a video[3], capturing a particular navigation through the interactive narrative, with audio explanations. We remark that both visual versions are implemented in French; the figures and data samples presented in this paper were translated for ensuring readability.

We used Tableau Desktop to render the interactive narrative and OBS Studio for recording the video.

Dashboard implementation We tested several visualisation possibilities, set up the various charts, maps and tables, and supplemented them with explanatory text, visual effects and audio explanations.

Each act is represented by an interactive dashboard, with the exception of Act VII (geography), which needing to display several maps, is broken down into several dashboards. Each episode (concerning a message) is represented by one or more visual artefacts (e.g. charts, maps, tables, text, images, audio).

As an example, Figure 4 presents the rendering of Act II (general presentation). It takes the form of a welcome

[3]https://youtu.be/u_KoBWc_qJU

screen, which gives access to the dashboards corresponding to the following acts (in the top menu). Several types of visualizations are used to highlight the messages. The interface offers the possibility of filtering by neighbourhood, district and year, for in-depth spatio-temporal focus, and to zoom in a map region.

4. Lessons learned

The crafting process is inspired by state-of-the-art models and processes [2, 6, 7]. However, several peculiarities of the application context have led us to enrich this process. This section presents the main lessons learned during this adaptation.

First, statistical data analysis is not sufficient for public health decision making. A systematic comparison with the state of the art, by comparing the figures obtained, is imperative in order to discern global phenomena from regional or seasonal peculiarities. Thus, decision-makers can judge which dimensions of the patient profile are in agreement with the situation in other countries, for which joint actions can be put in place, and which relate to the Gabonese population. Similarly, the results obtained must undergo extensive testing in order to prove its statistical value. As the target audience is predominantly scientific, these results can be communicated in the narrative.

Second, unlike seasonal narratives (frequent in data journalism), in scientific narratives analytical questions are not all known in advance. On the contrary, new questions may arise during data analysis. We illustrated this during the study of the lost to follow-up, in Subsection 3.2. Iterations between goal setting and data exploration phases are often necessary. New findings can also impact previous messages and require updating.

Third, the geographic component is very important in assessing the spatial and spatiotemporal extent of health problems. The restitution in the form of maps is to be privileged, but also, the spatial correlations.

Finally, the data narrative should allow interactive navigation between dashboards. There are different profiles among the decision-makers. On the one hand, we find various needs in terms of dimensions and indicators studied, for which a thematic organization (such as that implemented in acts III to VII) is perfectly suited. On the other hand, health authorities need a more comprehensive and guided reading of the narrative. The challenge is to find a good balance for rendering, both guided and interactive.

5. Conclusion

In this article, we have described the process of crafting a data narrative about TB pandemic in Gabon.

Unlike previous works, which target the general public, this narration targets experts in epidemiology and public health, who are decision-makers.

To our knowledge, this article is the first to describe the process of crafting a data narrative considering the peculiarities of the epidemiological field and highlighting the main methodological challenges. Concretely, (i) data analysis is completed by a comparison with other state-or-the-art studies, (ii) messages are supported by statistical tests, (iii) the geographic component is very important, (iv) the restitution must be both guided and interactive, and (v) back and forth transitions among process phases and steps are be necessary.

In addition to Gabonese health authorities, who are fully satisfied by the experience, the data narrative was presented to professionals of other African countries [27]. We hope that this initiative will serve to inspire other teams to reproduce the experience in other health fields, in particular to facilitate understanding of the epidemiological situation of other infectious diseases (covid, cholera, dysentery, yellow fever, ebola, etc.).

As perspectives, we plan to work on producing a data narrative based on the simulation results. Concretely, we are studying the spread of the tuberculosis epidemic with a multi-agent system whose parameters will be taken from the underlying spacial data warehouse, and which will be used to test several scenarios and health policies.

References

[1] J. Hullman, S. M. Drucker, N. H. Riche, B. Lee, D. Fisher, E. Adar, A deeper understanding of sequence in narrative visualization, IEEE TVCG 19 (2013) 2406–2415.

[2] F. E. Outa, M. Francia, P. Marcel, V. Peralta, P. Vassiliadis, Towards a conceptual model for data narratives, in: ER 2020, Vienna, Austria, 2020.

[3] R. Kosara, J. D. Mackinlay, Storytelling: The next step for visualization, IEEE Computer 46 (2013) 44–50.

[4] M. Bouman, Storytelling makes public health statistics more accessible, European Journal of Public Health 27 (2017).

[5] S. Rivest, Y. Bédard, M.-J. Proulx, M. Nadeau, F. Hubert, J. Pastor, Solap technology: Merging business intelligence with geospatial technology for interactive spatio-temporal exploration and analysis of data, ISPRS Journal of Photogrammetry and Remote Sensing 60 (2005) 17–33.

[6] S. Chen, J. Li, G. L. Andrienko, N. V. Andrienko, Y. Wang, P. H. Nguyen, C. Turkay, Supporting story synthesis: Bridging the gap between visual analytics and storytelling, IEEE Trans. Vis. Comput. Graph. 26 (2020) 2499–2516.

[7] B. Lee, N. H. Riche, P. Isenberg, S. Carpendale, More than telling a story: Transforming data into visually shared stories, IEEE Computer Graphics and Applications 35 (2015).

[8] F. E. Outa, M. Francia, P. Marcel, V. Peralta, P. Vassiliadis, Supporting the generation of data narratives, in: ER Forum, Demo and Posters 2020, Vienna, Austria, 2020.

[9] World Health Organization, Early detection, assessment and response to acute public health events: implementation of early warning and response with a focus on event-based surveillance, Technical document, 2014.

[10] N. Patel, N. Patel, Modern technology and its use as storytelling communication strategy in public health, MOJ Public Health 6 (2017) 338–341.

[11] B. McCall, L. Shallcross, M. Wilson, C. Fuller, A. Hayward, Storytelling as a research tool and intervention around public health perceptions and behaviour: a protocol for a systematic narrative review, BMJ open 9 (2019).

[12] E. K. Tsui, A. Starecheski, Uses of oral history and digital storytelling in public health research and practice, Public health 154 (2018) 24–30.

[13] L. K. Larkey, J. Gonzalez, Storytelling for promoting colorectal cancer prevention and early detection among latinos, Patient education and counseling 67 (2007) 272–278.

[14] B. Ezegbe, C. Eseadi, M. O. Ede, J. N. Igbo, A. Aneke, D. Mezieobi, G. C. Ugwu, A. U. Ugwoezuonu, E. Elizabeth, K. R. Ede, et al., Efficacy of rational emotive digital storytelling intervention on knowledge and risk perception of hiv/aids among schoolchildren in nigeria, Medicine 97 (2018).

[15] J. W. Njeru, C. A. Patten, M. M. Hanza, T. A. Brockman, J. L. Ridgeway, J. A. Weis, M. M. Clark, M. Goodson, A. Osman, G. Porraz-Capetillo, et al., Stories for change: development of a diabetes digital storytelling intervention for refugees and immigrants to minnesota using qualitative methods, BMC public health 15 (2015) 1–11.

[16] M. Cueva, R. Kuhnley, L. Revels, N. E. Schoenberg, M. Dignan, Digital storytelling: a tool for health promotion and cancer awareness in rural alaskan communities, International journal of circumpolar health 74 (2015).

[17] T. Botsis, J. E. Fairman, M. B. Moran, V. Anagnostou, Visual storytelling enhances knowledge dissemination in biomedical science, Journal of biomedical informatics 107 (2020).

[18] A. S. Peddireddy, D. Xie, P. Patil, M. L. Wilson, D. Machi, S. Venkatramanan, B. Klahn, P. Porebski, P. Bhattacharya, S. Dumbre, et al., From 5vs to 6cs: Operationalizing epidemic data management with covid-19 surveillance, in: 2020 IEEE Int. Conf. on Big Data, 2020, pp. 1380–1387.

[19] Y. Mboumba Sambo, Lutte contre la consommation de la drogue en milieu scolaire au gabon: Cas du Lycée Paul IDJENDJE GONDJOUT de la commune de Libreville, Dissertation, INJS Gabon, 2018.

[20] E. Engohan Alloghe, M. Toung Mve, S. Ramarojoana, J. J. Iba Ba, D. Nkoghe, Epidemiologie de tuberculose infantile au centre antituberculeux de libreville de 1997–2001, Med trop 66 (2006) 469–471.

[21] B. Melki, S. Saad, H. Daghfous, M. Khelifa, F. Tritar, Forme grave de la tuberculose : le pyopneumothorax tuberculeux, Revue des Maladies Respiratoires 32 (2015).

[22] B. Larbani, M. Terniche, S. Taright, M. Makhloufi, La prise en charge de la tuberculose pulmonaire dans une unité de contrôle de la tuberculose d'alger, Revue des Maladies Respiratoires 34 (2017).

[23] A. Sylla, B. Marchou, N. Kassi, N. Ello, T. Aba, G. Kouakou, C. Mossou, E. Ehui, S. Eholié, E. Biassagnéné, Co-infection tuberculose/vih: à propos de 717 cas suivis dans un service de maladies infectieuses en afrique subsaharienne, Médecine et Maladies Infectieuses 47 (2017) S137–S138.

[24] G. Tékpa, V. Fikouma, R. M. M. Téngothi, J. de Dieu Longo, A. P. A. Woyengba, B. Koffi, Aspects épidémiologiques et cliniques de la tuberculose en milieu hospitalier à bangui, The Pan African Medical Journal 33 (2019).

[25] S. NIANG, E. ABDALLAHI, K. THIAM, F. B. R. MBAYE, M. CISSE, A. DIENG, N. T. BADIANE, et al., Aspects épidémiologiques, diagnostiques et évolutifs de la tuberculose pulmonaire à microscopie positive au district sanitaire de saint-louis., Revue Africaine de Médecine Interne 5 (2018) 65–69.

[26] E. Segel, J. Heer, Narrative visualization: Telling stories with data, IEEE TVCG 16 (2010).

[27] R. Ondzigue Mbenga, V. Peralta, T. Devogele, S. Maghendji, E. B. Ngoungou, Processus de narration de données en intelligence épidémique avec application à la pandémie de tuberculose au gabon, in: 8e Journées Camerounaises d'Informatique Médicale, 2021.