Humane Visual AI: Telling the Stories Behind a Medical Condition

Wonyoung So, Edyta P. Bogucka, Sanja Šćepanović, Sagar Joglekar, Ke Zhou, and Daniele Quercia
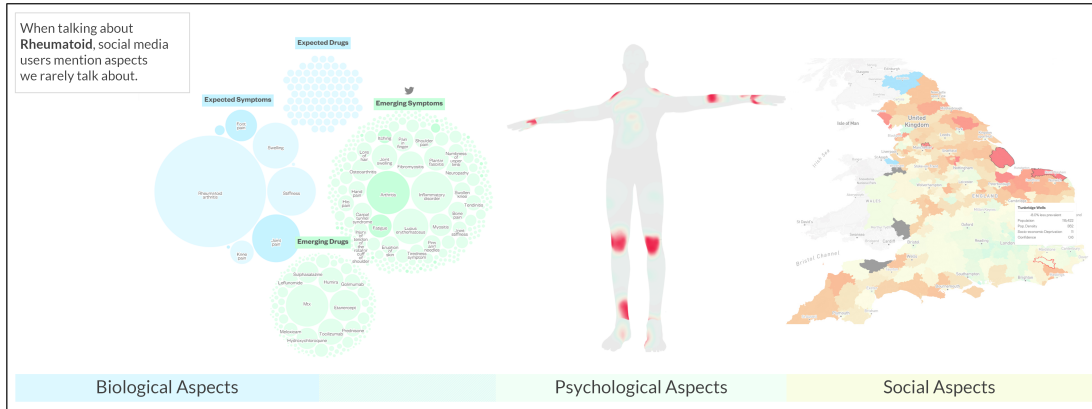
Fig. 1. Narrating the story behind a medical condition through biological, psychological, and socio-environmental lens (i.e., through the bio-psycho-social model).

Abstract— A biological understanding is key for managing medical conditions, yet psychological and social aspects matter too. The main problem is that these two aspects are hard to *quantify* and inherently difficult to *communicate*. To *quantify* psychological aspects, this work mined around half a million Reddit posts in the sub-communities specialised in 14 medical conditions, and it did so with a new deep-learning framework. In so doing, it was able to associate mentions of medical conditions with those of emotions. To then *quantify* social aspects, this work designed a probabilistic approach that mines open prescription data from the National Health Service in England to compute the prevalence of drug prescriptions, and to relate such a prevalence to census data. To finally *visually communicate* each medical condition's biological, psychological, and social aspects through storytelling, we designed a narrative-style layered Martini Glass visualization. In a user study involving 52 participants, after interacting with our visualization, a considerable number of them changed their mind on previously held opinions: 10% gave more importance to the psychological aspects of medical conditions, and 27% were more favourable to the use of social media data in healthcare, suggesting the importance of persuasive elements in interactive visualizations.

Index Terms—complex problem communication, storytelling, AI, social media data, healthcare, Martini Glass structure

◆

1 INTRODUCTION

The communication between patient and healthcare professional often focuses on the biological aspects of a medical condition (e.g., high temperature or pain) and overlooks its psychological and social aspects (e.g., patient feeling sad or lacking control due to recent difficult events in their social life). Yet these two latter aspects are considered to be equally important by the modern "bio-psycho-social model" in healthcare [21, 23, 59]. Social and psychological aspects such as fear, rage, neglect, attachment, mental well-being, life course, social capital, healthy lifestyle, neighborhood disadvantage, and perceived discrim-

ination are all shown to have important effects on chronic conditions such as obesity, coronary heart disease, and skin disease [21, 53], as well as on functional disorders [51, 61], and medically unexplained symptoms [21, 23]. After reviewing the literature, we found the most important psychological and social aspects to be the following: i) the patients' *emotions* (such as fear or trust) [15, 21], ii) their *psychological symptoms* (such as depression and anxiety) [4, 21], and iii) the *socio-economic status* of the areas where they live (such as neighbourhood disadvantage and social network) [2, 18, 21].

The main problem is that psychological and social aspects are: 1) hard to quantify; and 2) inherently difficult to communicate. First, they are hard to quantify because they have not been systematically recorded. To see why, consider the psychological aspects. There is no protocol in National Health Services that mandates the systematic recording of patients' feelings and emotions. Yet, in recent years, individuals have increasingly discussed their medical conditions on sites such as: *PatientsLikeMe* where they connect with others suffering from the same condition; *AskaPatient* where they share experiences with drugs; and the general-purpose *Reddit* and *Twitter*. It turns out that people tend to reveal even more personal experiences online than they would usually do in face-to-face interactions with their doctors (e.g., they discuss emotions, fears, and life events) – this is what has been called the disinhibition effect [58]. The problem with social media text is that it is unstructured and, as such, it is hard to process for the reliable extraction of medical conditions. The quantification of social aspects is challenging too – it is hard to relate socio-economic data (e.g., census data) to health outcomes.

- *Wonyoung So is with Massachusetts Institute of Technology, Cambridge, MA (US). E-mail: wso@mit.edu.*
- *Edyta P. Bogucka is with Chair of Cartography, Technical University of Munich (DE). E-mail: e.p.bogucka@tum.de.*
- *Sanja Šćepanović is Nokia Bell Labs, Cambridge (UK), E-mail: sanja.scepanovic@nokia-bell-labs.com.*
- *Sagar Joglekar is Nokia Bell Labs, Cambridge (UK), E-mail: sagar.joglekar@nokia-bell-labs.com.*
- *Ke Zhou is Nokia Bell Labs, Cambridge (UK), E-mail: ke.zhou@nokia-bell-labs.com.*
- *Daniele Quercia is Nokia Bell Labs, Cambridge (UK), CUSP King's College London. E-mail: daniele.quercia@nokia-bell-labs.com.*

Second, even if one were to be able to quantify psychological and social aspects, these would still be inherently difficult to communicate. That is simply because these aspects are part of a multi-faceted bio-psycho-social model. Multi-dimensional data is hard to visualize in engaging ways [28], and this represents an active research area [35, 36, 54]. Visual storytelling was suggested to be the next-level approach in healthcare communication [8, 40].

This work tackled the two challenges of both *quantifying* and *visually communicating* a medical condition's psychological and social aspects. In so doing, it made three main contributions:

C1: By considering around half a million Reddit posts in the sub-communities specialised in 14 medical conditions, for each condition, we *quantified* its psychological and social aspects by using (Section 3): a) our new deep-learning framework to extract medical entities from the posts with state-of-the-art accuracy [63]; b) a widely adopted dictionary-based approach to extract emotions from the posts; and c) a probabilistic approach that mines open prescription data from the National Health Service in England to compute the prevalence of drug prescriptions related to the condition.

C2: We *visually communicated* each medical condition's biological, psychological, and social aspects through a layered Martini Glass visualization (Section 4).

C3: We evaluated our visualization and found that it had persuasive effects (Section 5). In a user study, after interacting with our visualization, a considerable number of our 52 participants changed their mind: 10% gave more importance to the psychological and social aspects of medical conditions; and 27% were more favourable to the use of social media data, and 10% to the use of AI in healthcare, suggesting the importance of persuasive elements in interactive visualizations.

The resulting visualization is available on the project's page `http://humane-ai.social-dynamics.net`.

2 RELATED WORK

There have been two lines of past research related to ours. The first concerns general ways of quantifying health outcomes. The second concerns visualizing multi-faceted health aspects.

General ways of quantifying health outcomes. One of the primary tasks in mining health discussions is to extract medical mentions. Initial approaches to this task were *keyword and lexicons-based* [25, 32, 47]. They worked reasonably well on formal medical text but had well-known limitations when applied to social media text: they failed to capture the variability of informal language, diversity of users' symptom expressions, and spelling mistakes [13, 49]. The approaches that dealt better with such limitations were based on *machine learning*, such as Conditional Random Fields (CRFs) [34, 43] and on *deep learning*, such as Recurrent Neural Networks (RNNs) [60, 62]. *Dictionary-based approaches* were used to extract linguistic features from posts, which were then shown to relate to chronic disease health outcomes, both on Twitter [16] and Reddit [12]. Psychological and social aspects (such as positive emotions, work-related issues, or stigma) were studied on Reddit [48] and on Twitter [55]. The drug prescription data from the general practitioners across England were used as a *proxy for prevalence of diseases*: examples of such uses include the health domains of diabetes [2], influenza [50], and opioid addiction [17].

Visualizing multi-faceted information: Dashboard and Narrative Visualization. Dashboards have been one of the most popular choices for visualizing medical data, as they can consolidate multi-faceted clinical conditions in one screen [20, 22, 31]. However, clinical dashboards mainly supported medical professionals [29], and, more generally, dashboard visualizations have been found to have two main weaknesses: information overload given the single screen [65], and a significant learning curve [29].

An alternative visualization technique for delivering complex information is the *narrative visualization*. The idea is to create a story based on the results of the data analysis. This technique has been widely used
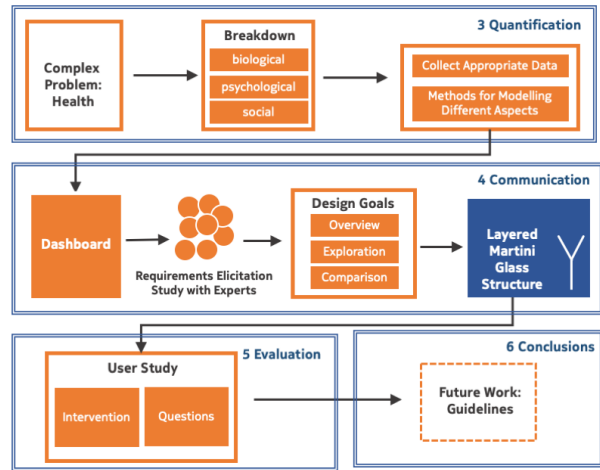

Fig. 2. Steps to visualize the complexity behind health, and evaluate the resulting visualization. After breaking down the problem into its three main aspects, we quantified each of these aspects by collecting appropriate data upon which we run state-of-the-art machine learning models *(top row)*. We determined our design goals by conducting a requirements elicitation study with medical experts, and then built a layered Martini Glass Structure visualization *(mid row)*. We conducted a user study upon which we drew guidelines for future reference *(bottom row)*.

by several news organizations such as the New York Times and The Guardian, and its effectiveness has been extensively studied [6, 7, 9]. Segel and Heer [57] highlighted the importance of balancing between author-driven parts of a narrative visualization, usually linear and non-interactive, and reader-driven ones, usually interactive and exploratory. Hullman and Diakopoulos [27] highlighted the importance of "rhetorical" and "persuasive" elements when communicating a story in visualizations. Drawing on both seminal papers, Lee et al. [33] further articulated visual storytelling process starting from data exploration to story delivery to users' feedback.

In summary, previous work on *quantifying* health outcomes has not been sufficiently general, in that, it has focused on very specific medical conditions or symptoms, and it has done so by focusing on one condition/symptom at a time. Also, previous work on *visually communicating* multi-faceted information in healthcare mainly targeted medical practitioners, and used dashboard-type of visualisations.

3 QUANTIFYING THE BIO-PSYCHO-SOCIAL MODEL

We resorted to Reddit, an online community in which people exchange ideas and provide support to each other. Subcommunities in Reddit called subreddits consist of subsidiary threads that allow users to focus on a specific interest or topic. When deciding which subreddits to gather, we focused on the conditions with distinctive symptoms, making them easier to discriminate from others (e.g., psoriasis) and excluded complex and hard to discriminate conditions (e.g., cancer). Based on the data availability, we incorporated the following 14 conditions: rheumatoid arthritis, diabetes, depression, dementia, borderline personality disorder, psoriasis, gastroparesis, hypothyroidism, irritable bowel syndrome, dementia, Parkinson's disease, Meniere's disease, multiple sclerosis, and sleep apnea.

We collated all the Reddit posts in those subreddits during the first half of 2017, which resulted in about half a million posts in total.

3.1 Quantifying Biological Aspects

The biological side of a medical condition is partly captured by symptoms and drugs associated with it.

Expected Symptoms and Drugs from Literature. We used MedScape, WebMed, and Wikipedia. From MedScape, a website providing information to physicians and health care professionals, for

a given condition, we obtained the list of signs and symptoms that are used to diagnose it. However, MedScape does not provide all the symptoms (i.e., it does not include the most common and generally non-discriminating symptoms, such as pain and inflammation). For this reason, we augmented this list of primary symptoms by manually searching WebMed and Wikipedia for additional ones.

Since the same symptom/drug can be expressed in several ways, doctors adopt a common technical terminology. This work adopted the most comprehensive and widely used terminology, which is called the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT). It is used in over 50 countries, and contains synonym expressions for each medical concept (for example, for *rheumatic arthritis*, synonyms include *rheumatic gout* and *proliferate arthritis*). We manually linked each of the previously collected symptoms/drugs to the corresponding concept in the SNOMED CT database. For instance, consider the symptom *weakness*. Using SNOMED CT's search mechanism, we linked it to the concept *asthenia* with the ID *13791008*. As a result, we associated a set of *expected symptoms and drugs* expressed in a proper medical language with each of the 14 medical conditions. For example, *rheumatoid arthritis* was associated with the symptoms of *stiffness, tenderness, swelling*, and *pain in joints*, and with the drugs *Leflunomide* and *Celecoxib*, among others.

3.2 Quantifying Psychological Aspects

In addition to the biological aspects, a medical condition is characterized by its psychological aspects, i.e., a rich set of individual experiences. To quantify such experiences, we needed to gather: i) *all the symptoms/drugs mentioned by different patients* (a significantly larger set of symptoms/drugs emerges than the expected ones), as well as ii) *the distinct emotions and body parts the patients mentioned*.

Emerging Symptoms and Drugs Online. From the Reddit dataset, we extracted symptom and drug mentions. Extracting medical mentions from free-form text is a challenging NLP task. People use different forms to express the same type of symptom (for instance, *brain fog* and *clouded consciousness*); use slang for medication names (for instance, *benzos* for *benzodiazepines*); or make spelling mistakes. An equally challenging task is to map the extracted mentions to an official terminology, such as SNOMED CT. For example, one needs a method to link both *brain fog* and *clouded consciousness* to the same concept, which happens to have the ID *40917007* in SNOMED CT.

Thanks to the advances in modern deep-learning techniques in Natural Language Processing, it has become possible to do both tasks. We implemented a state-of-the-art *entity extraction method* (sketched in Figure 3, detailed in Supplemental Material, and published in [62]) based on bidirectional long short-term memory (BiLSTM) units in combination with conditional random fields (CRFs) that utilizes the contextual string embeddings (i.e., RoBERTa embeddings) [38].

The symptoms/drugs mentioned on Reddit that were not among the expected ones were called *emerging symptoms/drugs*. To see why, consider that, for *rheumatoid arthritis*, in addition to its expected symptoms, a myriad of additional symptoms had emerged online: *hair loss, itching, exhaustion, bone pain, muscle weakness, mental dullness, heavy legs, tiredness,* and *bedridden*, to name a few. Such emerging symptoms are not discriminative of the condition itself but, instead, they reveal the diversity of the individual experiences among patients, and part of them represent the psychological aspects of *rheumatoid arthritis*.

Emotion Words. One of the primary forms of people's psychological experience is represented by emotions. To categorize emotions extracted from our social media data, we employed the Plutchik's psychoevolutionary theory [52]. The theory lists 8 basic emotions: anger, joy, disgust, fear, anticipation, trust, sadness, and surprise. We looked at how many words expressing each of these basic emotions were mentioned when talking about each of the 14 conditions. To determine whether a word belonged to an emotion category, we resorted to the "EmoLex" word-emotion lexicon [41], which classifies words into the eight categories.
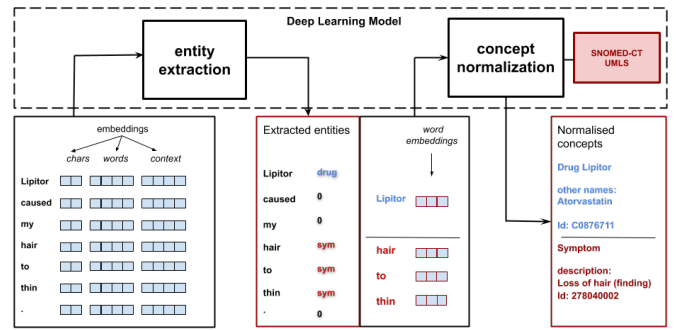

Fig. 3. Sketch of our deep learning models that extract medical mentions from social media, and associate them with concepts present in our medical hierarchy (SNOMED CT).

For each word (e.g., "cry","tear") related to a specific emotion category (e.g., "sadness"), and a condition (e.g., "depression"), we computed the Point-wise Mutual Information (PMI). Our PMI analysis (detailed in Supplemental Material) compares the observed occurrence probability to the chance probability of occurrence given a uni-gram model of independent interactions. It considers that words relevant to a keyword will occur with it far more often than would be expected by chance.

By aggregating all the words ("sad", "tear", "cry") belonging to a particular emotion category ("sadness"), we computed the association between each condition d and each emotion category cat.

Body Parts. There is previous research linking experience of emotions to different body parts [44], so our visualization showed which body parts people mentioned with each of the 14 conditions. To determine the associations between body parts (e.g., "mouth") and a given medical conditions, we applied a methodology similar to the previous one used for emotions (i.e., a TF-IDF normalization). To that end, we created a 'body' dictionary that lists all the words (e.g., "tongue", "lip") related to a specific body part (e.g.,"mouth") by combining the original LIWC dictionary subcategory "body" and the Wikipedia "human body and organs."[1]

3.3 Quantifying Social Aspects

People's social environment can cause or affect their health. For example, the economic status has an influence on people's diets, and this, in turn, can affect whether they will suffer from diabetes [2]. On the other hand, people's medical condition can influence their social life – for instance, cancer patients experiencing severe pain are likely to suffer from lowered-self esteem and anxiety [45]. To associate socio-economic conditions with disease prevalence, we associated census data with prescription data across all the boroughs in England.

Prescription data is released by the National Health Service (NHS) and enumerates all the prescriptions done by any general practice (GP) across England. The drug prescription data has been made available since July 2010, is released every month by NHS [1], and consists of four files (Figure S2 the Supplemental Material), which we collected until the end of June 2019: i) *NHS GP monthly prescriptions:* each prescription contains the drug's name, its British National Formulary (BNF) code, the practice code, its total number of items, total cost, and each item's quantity; ii) *NHS drugs:* each row contains a drug (individual preparation name) and its unique BNF Code; iii) *NHS GPs:* each row contains a practice code, name, and full address; and iv) *NHS GP patients:* each row contains a practice code, census LSOA borough code (each row refers to a specific (practice, LSOA) pair), and the number of the practice's patients in that borough. From this file, we computed the total number of primary care patients who live in each borough.

---

[1] https://en.wikipedia.org/wiki/List_of_organs_of_the_human_body

The data, however, does not contain any mapping between the prescribed drugs and the corresponding medical conditions. To create such a mapping, we crawled the entire database from DrugBank[2]. From the database, we obtained 9105 drug-related pages. Out of those, we were able to match 3013 to at least one associated condition. This gave us a mapping between drugs and conditions. More specifically, for a given condition C, we used DrugBank to find the set of drugs D associated with it to then estimate the total prescriptions related to C in a borough A.

4 COMMUNICATING THE BIO-PSYCHO-SOCIAL MODEL

We first developed a visualization for medical professionals, with whom we conducted a requirements elicitation study (Section 4.1). To fulfill the design goals that emerged from this study, we then surveyed and selected the most appropriate visualization approaches (Section 4.2). These approaches were then used to develop the final visualization aimed at the general public (Section 4.3).

4.1 Communicating with a Dashboard

The purpose of our study with medical professionals was twofold. First, we wanted them to assess the medical content and validity of the visualization. Second, as doctors convey medical knowledge to patients on a daily basis, we relied on their experience and asked them if the general public could be interested in such a visualization, and if so, what functionalities would be of help. From their answers, we selected several requirements and design suggestions.

We found that the medical experts were already aware of the bio-psycho-social model, and were familiar with dashboards [5, 14]. Hence, for our elicitation study, we developed a dashboard prototype.

Since dashboards aim at showing all the information at once, we provided the functionality to select a medical condition, for which we then showed its real-world prevalence (left panel in Figure 4) and associated symptoms (right panel in Figure 4). When a user clicked on a region on the map, then the dashboard showed detailed information about the region, including the prevalence of the condition, the confidence value, and its socio-economic indicators such as deprivation index, population, and population density. The two sides of the dashboard were interactively linked; when a user clicked on a symptom bubble, then the choropleth map would change to display its prevalence.

The goal of the elicitation study was to discover the main requirements for visualizing the bio-psycho-social model.

We recruited a group of 13 medical experts, who interacted with the dashboard and answered an online survey afterwards. Our domain experts were based in the United Kingdom (4 experts), United States (2), Italy (2) Ecuador (1), Kenya (1), Spain (1), and 2 preferred not to disclose their location; 3 of them were epidemiologists, 3 were Public Health researchers, 2 were clinical academics, 1 was a Public Health manager, while 4 preferred not to provide this information.

We asked them three open-ended questions:

Q1 How would you describe the *intuitiveness* of this tool?
Q2 Could the general public be *interested* in such a visualization and, if so, why?
Q3 Which *additional features* or functionalities would you like to see in a tool aimed at the general public?

We then summarized their replies to each question.

[Q1] Intuitiveness.   In general, most of the experts found the tool to be intuitive. However, they also expressed difficulty with having the different aspects of a medical condition collated into a single screen. They mentioned three main shortcomings: i) *lack of self-explanation* (without a proper tutorial, it was hard to use the visualization to its full potential), ii) *unclear links among data sources* (the respondents perceived that the combination of multiple datasets was not visually apparent, e.g., the relationship between the prescription map and the symptoms extracted), and iii) *screen real estate issue* (i.e., the space is too limited for communicating complex information).

[Q2] Interest of the General Public.   They found the tool could be of interest to general public for three main reasons: curiosity and novelty, a better understanding of the symptom-condition relationships, and ability to discover their areas' health.

In terms of *curiosity* they said: *"People might like to play around with such visualizations out of curiosity"*, *"Absolutely - curiosity"*, *"Yes. If done right, it can spark curiosity like ("oh I have depression, where are people similar to me?")"*. One expert pointed out the *novelty* of the tool as a factor that could spark interest: *"The general public would be interested for the novelty factor as there is no such tool that can bring these dimensions together."*

Another aspect they mentioned was the ability to *better understand the relationships between symptoms and medical conditions*. For instance, professionals said that the tool *"provides rapid information on symptoms associated with a specific disease and therapy"*, and *"[...] it is a powerful tool to explore symptoms for lots of diseases."* As *"many diseases and conditions have symptoms that are uncommon and rare, doctors tend to dismiss these symptoms, leaving the patients confused and scared."* By using the tool, the general public can *"feel how common is the disease and get a sense of companionship, if the set of signs and symptoms they present match a known disease."* Finally, some of the experts said: *"Also, it could be a good way to communicate news about health."*, and *"People are increasingly interested and wanting to be involved, share and contribute to health issues."*

The third group of comments had to do with the *health status of areas*. For instance, they mentioned: *"Yes, because of the map: people would like to see what are the symptoms/diseases that are more common in the area they live in."*; or *"To have a general idea about diseases affecting their area."*, or *"I think they would be interested since they would like to know what the current disease outbreaks in neighborhoods so that they can prepare in advance."*, and one expert went further to state *"...figuring out the best places to live/visit, if the socio-economic indicators are incorporated."*

[Q3] Additional Features.   The experts mentioned six main additional features: i) providing *a tutorial* to guide the users through the complexity of the model and the visualization, ii) offering *background* information about the results calculation, iii) highlighting *uncommon symptoms*, iv) *rankings of regions* according to the prevalence of medical conditions, v) *improved labeling* of different parts of the visualization, and vi) *adding mouse hovering interactions*.

Design Goals.   After summarizing the requirements emerging from the elicitation study, we identified three main design goals:

Simplify user interaction and model exploration. The final tool should guide interaction and exploration instead of asking users to make complex choices.

Support uncommon symptoms-condition exploration. The tool should visually differentiate symptoms that are common to a condition from those that are uncommon.

Enable spatial exploration. The tool should provide users with means of selecting several areas and comparing them in terms of different indicators.

4.2 Communicating the Complexity

To meet the three design goals, by reviewing the literature, we adopted and identified three techniques: story synthesis, narrative visualization, and layering with animation.

First, we adopted *story synthesis* [11]. The idea is to select and organize the quantitative findings. Second, to show the organized findings, we adopted a *narrative visualization* approach [57]. In particular, we adopted the *Martini Glass Structure* identified by Segel and Heer [57], where the story begins with an explanation or narrative ("author-driven" part) and continues with a more exploratory and interactive data visualization ("reader-driven" part). Creating a linear-sequence would give us much more freedom than the dashboard because we can address each concept sequentially ("bio," "psycho," and "social"). Moreover, this Martini Glass-structured visualization could address the need of a tutorial without explicitly having it, as it can naturally increase the
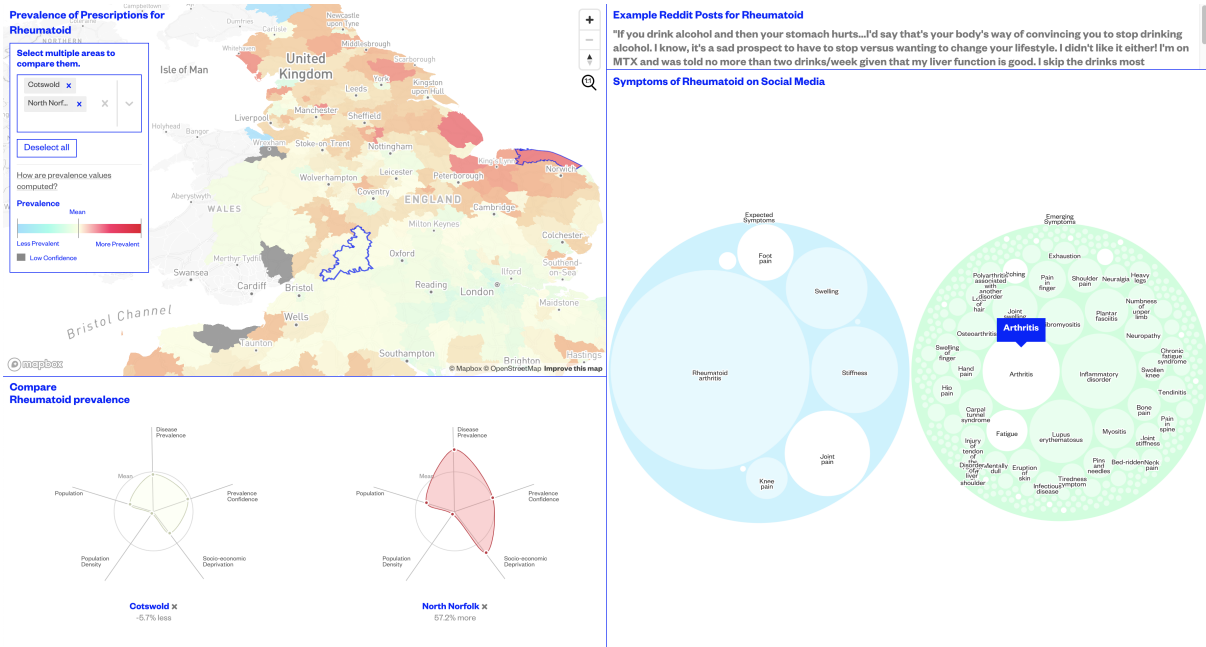
Fig. 4. The dashboard used in the elicitation study among medical experts. A map encodes prescription prevalence together with socio-economic indicators *(left)*, and blue/green bubbles show not only the typical symptoms but also those emerging from social media conversations *(right)*.

level of understanding while users go through the visualization. Third, we adopted the *layering technique* [56] to reduce the complexity of the bio-psycho-social model with its multiple points of view. The layering technique relies on breaking down the main point into multiple digestible chunks, and showing one chunk at a time. Also, drawing on Gestalt principles [56], several types of animations were incorporated (e.g., selection and scroll). The effectiveness of animations lies in guiding the users' attention to focus on the visualization's intended message. Such functionality allows the users to discover complex stories at their own pace.

To sum up, our design goals were met by a combination of three techniques: story synthesis, narrative visualization, and layering with animations. The *Martini Glass structure* enabled a reader-driven approach (the user's selections determined what content was shown). To avoid complex user interactions, we incrementally introduced the three components of the bio-psycho-social model with *layering* and *scrolly-telling techniques*.

4.3 Communicating with Layered Martini Glasses

We selected a suitable design representation for each of the three story slices: bubbles for the biological slice, human body and emoticons for the psychological, and a choropleth map for the social. Then we applied a layered Martini Glass structure (Figure 5) for the arrangement of the story slices. The user interaction between the layers is provided by linear scrolly-telling and animations. The first and last layers of the model are structured as two Martini Glasses. They start with our predefined author-driven narrative, and end up with a reader-driven exploration. Within each exploratory snippet, more complex interactions such as clicking, dragging, zooming, and panning are provided to offer personalized data views. The visualization has then multiple opportunities for users to share their "data views" using social media.

4.3.1 Introducing Biological Aspects

The visualization starts with a brief introduction of traditional biomedical model. The introduction, in turn, starts with the controversial statement that this biomedical model "doesn't fit the reality." This statement aims at capturing user's attention. Then, as a first interaction step, the Martini Glass structure allows the user to select a medical condition (e.g., rheumatoid arthritis) (Figure 5a).

Since traditional medical practice has focused on the biological side, the visualization starts with describing this layer by showing the *expected* associated symptoms and drugs (Section 3.1) for the selected condition, e.g., rheumatoid arthritis. The symptoms and drug names are represented as two blue Packed Bubble Charts (Figure 5b and Figure 6, left). The sizes of the bubbles vary based on how common the symptoms/drugs are, and is calculated using a TF-IDF normalization. Once the user hovers over a bubble, the number and the list of all the associated conditions are displayed. We chose to use Packed Bubble Chart [64] because of three main reasons. First, this type of diagram visualizes medical conditions as cohesive collections of symptoms/drugs with its aesthetically pleasing and almost organic appearance [42]. Second, it visually encodes the frequency of symptoms/drugs with circle sizes that enable value comparisons [26]. Third, it enables easy discovery of unanticipated patterns of uncommon and overlapping symptoms/drugs.

We kept the blank space on the right hand side of the blue bubbles (Figure 5b) to evoke tension and nudge the user to keep on scrolling down.

4.3.2 Adding Psychological Aspects

The visualization of the second layer (Figure 5c) starts with the green Packed Bubble Charts (Figure 6, middle). The green clusters show the symptoms and drug names *emerging* from social media. The animated transition from the expected symptoms and drug names to the emerging ones aims at softening the separation between these two types and at evoking surprise [56]. By comparing these two bubble clusters, the user is able to acknowledge that social media data can uncover unforeseen symptoms and drug names. The "share button" allows the user to post his/her current "data view" on Twitter.

Subsequently, the user is presented with emotions and associated body parts (Figure 5d and Figure 6, right). We tried to address the problem of intuitiveness previously emerged in relation to the dashboard and the need of showing emotional data with a direct visual encoding: (emoji-like) icons and a human body. The first visualization introduces the ranking of the 7 emotions in terms of how often they are mentioned with rheumatoid arthritis, e.g., "sadness", "disgust", "anger", "surprise", "fear", "trust," and "anticipation." We used emojis to represent the emotions because of their common use [19, 39]. Below the emotion ranking, the user can see and rotate a 3D-model of a human body with a
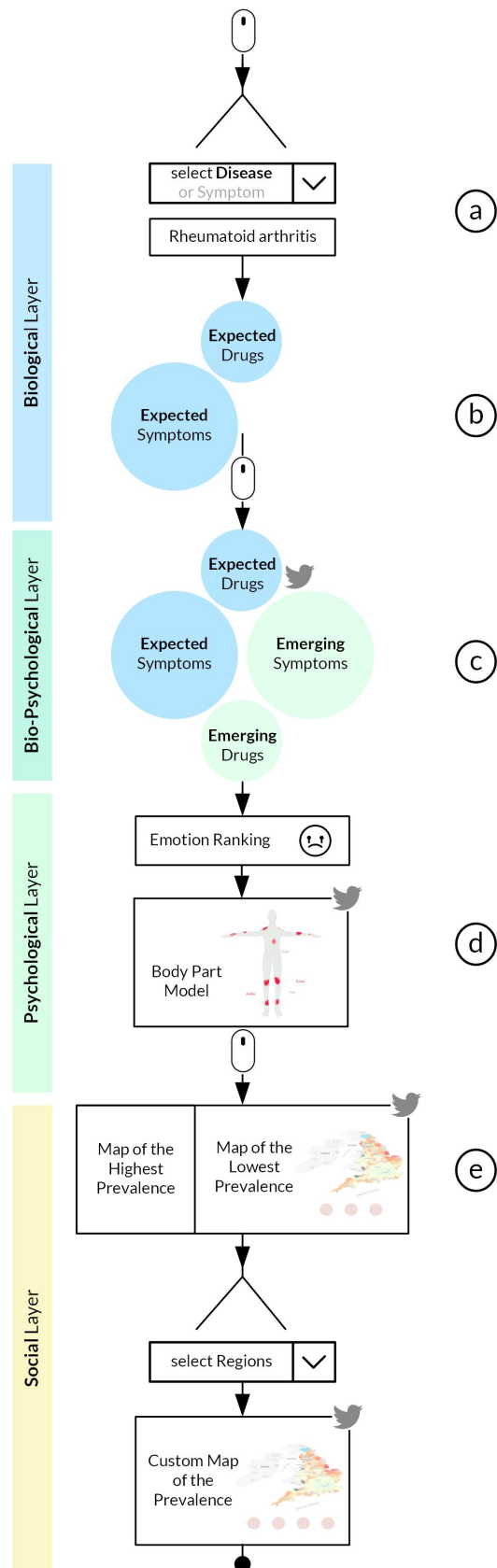
Fig. 5. The storyline of the visualization. Upon selecting a medical condition (a), the three layers–the biological (b); (a view shown in Figure 6, left), psychological (c and d); (Figure 6, mid and right), and social (e); (Figure 7)–are linearly stacked with two Martini Glass structures (one in the biological layer, and the other in the social layer). Each layer offers opportunities to share parts of the visualization on social media.

heat-map texture projected onto it. The texture illustrates specific body parts that social media users often mention when discussing the medical condition under consideration. We decided to mark this value with a texture on the surface of the body, instead of pointing at the exact body parts. This reduces the complexity of the visualization, and illustrates the inexact nature of social media expressions. The color and size of the body textures encode the magnitude of mentions. Our choice of including the body parts visualisation under the psychological aspects was motivated by previous work on bodily maps of emotions [44], which demonstrated that different emotions are associated with topographically recognizable bodily sensations. Again, the self-contained visualization can be fully shared by the user on Twitter.

4.3.3 Adding Social Aspects

The final layer (Figure 5e) contains a choropleth map of England and a series of radar charts showing a region's population, population density, and socio-economic deprivation (Figure 7, top). The diverging color scheme shows the prevalence of the prescriptions associated with the condition in relation to the mean value for England. The layer follows a Martini Glass structure. It first starts with an author-driven narrative, showing the regions with the highest and the lowest prevalence of prescriptions for the condition, and ends with interactive maps. The user can select up to four regions by choosing them from the list, or by clicking them on the map. Then the radar charts are accordingly updated, so the user can compare the prevalences across the four regions (Figure 7, bottom). For instance, the two regions with the highest rheumatoid arthritis prevalence are North Norfolk and Rochdale. The visualization shows that their socio-economic characteristics are almost the opposite: population size and population density of North Norfolk are way below the mean, while those of Rochdale are above the mean; and socio-economic deprivation is slightly above the mean for North Norfolk, while it is extremely high in Rochdale. A radar chart, also known as star plot or polar chart [10, 24], is a graphical method effective in visualizing multidimensional data with an arbitrary number of variables [3,30,37]. We selected the radar chart for two main reasons: *i)* it allows for easy comparison of ordinal values of socio-economic indicators across regions, and *ii)* it is suitable to convey outliers, i.e., the regions that have a value that is significantly above or below the country's average. To further increase user engagement, again, the comparative maps can be shared on Twitter.

5 EVALUATION

The goal of the study was to evaluate how well our visualization communicated the complex bio-psycho-social model to an audience that might not be familiar with it.

5.1 User Study Setup

The study consisted of four steps, taken from a method that was previously used for measuring the persuasive effects of data visualizations [46].

Step 1: Survey. Once the participants agreed to take part in the experiment, we asked them to respond to a 6-statement survey that assessed their opinions on how doctors should operate. Additionally, we asked them if they searched for health advices on the Internet and, if so, on which websites. Finally, we asked them to report their gender and age.

Step 2: Intervention. We asked the participants to interact with our visualization (the treatment) or with the baseline (the control) and to carefully investigate its content. This step relied on an attention-check question: we asked the treatment group to name two of the emotions associated with a medical condition and we asked the control group to name external websites the baseline visualization was referring to. Answers to either question could be known only if the participants paid attention to the visualization.

Step 3: Survey. After interacting with the treatment or control, participants were asked again to respond to the survey. This allowed us to determine any potential opinion change due to the treatment/control.
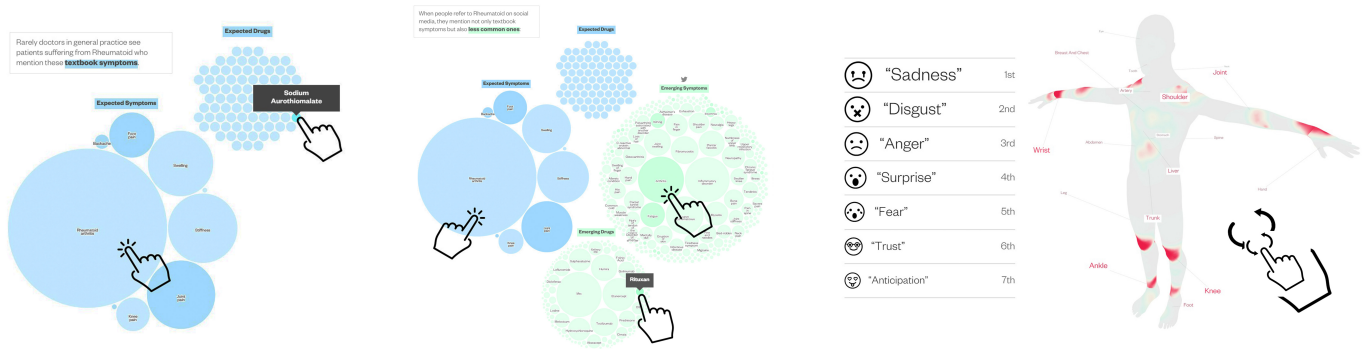
Fig. 6. Visualizing the biological and psychological aspects. The packed circle charts display expected (left, blue bubbles) and emerging (middle, green bubbles) symptoms for *rheumatoid arthritis*, together with corresponding drug names. The emotions associated with the condition and respective body parts are also shown (right).
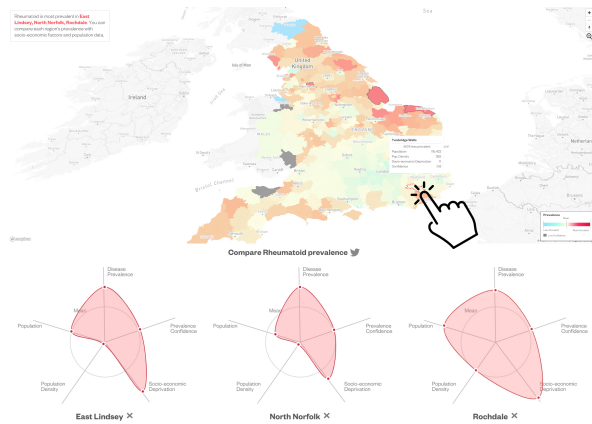

Fig. 7. Visualizing the social aspects with: a map of the prevalence of prescriptions for *rheumatoid arthritis (top)* (by clicking on a region, a pop-up window with its socio-economic indicators is shown); and a comparison of the three regions with the highest prescription prevalence of the condition across England *(bottom)*.

Step 4: Open-ended questions. We also asked the participants to elaborate on why they changed or did not change their opinion. For users interacting with the baseline, we also asked what charts or graphs should be added to the page to better explain symptoms of conditions, people's emotions, and their social circumstances. Participants could finally share any additional feedback.

Survey.   The survey consisted of six opinion determination statements. The first four statements started with the phrase *"In treating medical conditions, doctors should..."* and ended as follows:

... treat only the corresponding symptoms (S1). This statement aimed at assessing the acceptance level of the current biomedical model.
... consider also the emotional state of the patients (S2). This statement assessed the understanding and acceptance of the psychological aspects of conditions.
... consider also the social and economic backgrounds of the patients (S3). The statement assessed the understanding and acceptance of the social aspects of conditions.
... analyse what their patients shared on social media (S4). The statement revealed opinions towards using social media data in healthcare (impersonal perspective).

The last two statements were, instead, self-contained:
In treating your medical conditions, your doctor should analyse what you shared on social media (S5). This statement assessed whether the participants would share their own social media data for health purposes (personal perspective).

Artificial Intelligence can improve healthcare (S6). The statement aimed at capturing their general opinions on the matter.

The participants were asked to consider each statement and decide to what extent they agreed with it on a 7-point Likert scale, ranging from "strongly disagree" (-3) to "strongly agree" (+3).

The experiment was conducted using the Amazon Mechanical Turk platform. The recruited participants fulfilled the following criteria: (i) self-reported location in the UK (to be able to relate to geographic prescription prevalence, which is shown on a UK map or described in the baseline); (ii) previous task approval rate of at least 95%; and (iii) number of previously approved tasks higher than 5000. Each experiment took between 5-10 minutes, and participants were paid 1 USD for completing it.

Baseline.   The baseline for the control group (shown in Supplemental Material) had the following narrative structure: a short explanation of the selected condition extracted from Wikipedia; three pre-selected posts from Reddit mentioning symptoms, drug names, and emotions; and a table showing the prevalence of the condition and associated socio-economic factors.

5.2 Evaluation Metrics

To measure potential opinion changes, one of the commonly used approaches is to segment the participants into three *opinion groups* based on their answers before and after exposure to the treatment/control [46]:

Negatively Polarized (NP): *strongly disagree* (-3), and *disagree* (-2);

Neutral/Weakly Polarized (NWP): *somewhat disagree* (-1), *neither agree nor disagree* (0), and *somewhat agree* (+1);

Positively Polarized (PP): *agree* (+2), and *strongly agree* (+3).

By using these three groups, we calculated two metrics for each of the six statements. First, we calculated the *persuasion likelihood* as the percentage growth rate of each opinion category (ΔNP, ΔNWP, and ΔPP) after the user exposure to the visualization [46]. For example, the positively polarized category $\Delta PP = (PP_{after} - PP_{before})$, where PP_{after} is the percentage of participants who were positively polarized towards a statement *after* experiencing the visualization, and PP_{before} is the percentage of participants who were so even *before* experiencing it. In a similar way, we computed the persuasive likelihood across the two remaining opinion categories: $\Delta NWP = (NWP_{after} - NWP_{before})$, and $\Delta NP = (NP_{after} - NP_{before})$. Second, we calculated the *average opinion change* per statement, which is the mean across all participants of these Δs.

5.3 Results

The study was completed by 108 participants: 52 of them interacted with our visualization, and 56 with the baseline. The set of participants was diverse in terms of age (18-64 years, median age was 30), and gender (38% females, 58% males, and 4% of the users preferred not to answer this question).

5.3.1 Opinions Before the Intervention (Step 1)

Table 1 summarizes the users' *initial* opinions. The opinions on the current biological model were almost equally distributed among the different groups. Considering the three aspects of the bio-psycho-social model, the most positive opinions were towards its psychological aspect: 63% of all participants agreed that doctors should consider the emotional state of their patients. The idea of including social and economic backgrounds tended to elicit either neutral responses (44%) or positive ones (42%). Unsurprisingly, participants were considerably against the general idea of sharing social media data with doctors (65%). This opinion got even more polarized once they were asked to think about their own data (71% of negative answers). The majority of the participants already had a neutral (46%) or positive opinion (50%) towards the use of AI in healthcare.

To search for reliable health advises our participants usually browsed professional medical websites such as the official portal of NHS (40 mentions) and WebMD (17 mentions). The trend of using social media as a source of health-related information was weak among them – only 3 mentioned the use of Reddit, but 24 said they performed general health searches on Google.

#	Statement	NP	NWP	PP
In treating medical conditions, doctors should:				
S1	treat only the corresponding symptoms	.31	**.37**	.33
S2	consider also the emotional state of patients	.04	.33	**.63**
S3	consider also the socio-economic background of patients	.13	**.44**	.42
S4	analyse what their patients shared on social media	**.65**	.23	.12
In treating your medial conditions, your doctor should:				
S5	analyse what you shared on social media	**.71**	.19	.10
Artificial Intelligence can:				
S6	improve healthcare	.04	.46	**.50**

Table 1. Distribution of the participants' initial opinions per statement. The percentages of responses in each of the negatively polarized (NP), negatively/weakly polarized (NWP), or positively polarized (PP) categories are shown. Participants' opinions were equally distributed across the categories for the biological aspect (S1), tended to be positive about the psychological aspect (S2), and were split between positive and neutral views about the sociological aspect (S3). When it came to the use of social media data in healthcare, our participants were negatively polarized (S4 and S5). The use of Artificial Intelligence was mostly viewed as being either positive or neutral.

5.3.2 Opinions After the Intervention (Step 3)

To test for statistical significance of the observed patterns, we used the Wilcoxon Signed-Rank Test, which tested the null hypothesis that the exposure to the treatment or control did not affect a participant's opinion. The findings were considered statistically significant at the level of $p < .05$.

Persuasion Likelihood. Table 2 shows the persuasion likelihood by statement after being exposed to our visualization. The participants tended to slightly decrease their support for the biological model (S1). Those who were neutral about the importance of psychological aspects switched their opinion to a more positive view (Δ 10%) (S2), while those neutral about the importance of socio-economic conditions took either a more negative or more positive stance and, as such, became more polarized (S3).

Those who were against the use of social media switched to stances that were more neutral (21%) or more positive (6%) (S4 and S5). Finally, 10% switched to a more positive view of the general role of Artificial Intelligence in healthcare (S6).

Average Opinion Change. Figure 8 shows the average opinion change by statement at a 95% confidence. Changes were statistically significant for the last three statements for both baseline and our visualization. The highest changes concerned the sharing of the social media data with doctors to improve treatment possibilities (S4 and S5). Overall, the persuasive effect of the visualization was stronger than that of the baseline.

#	Statement	NP	NWP	PP
In treating medical conditions, doctors should:				
S1	treat only the corresponding symptoms	.00	**.04**	-.04
S2	consider also the emotional state of patients	.00	-.10	**.10**
S3	consider also the socio-economic background of patients	.06	-.12	.06
S4	analyse what their patients shared on social media	-.27	**.21**	.06
In treating your medial conditions, your doctor should:				
S5	analyse what you shared on social media	-.27	**.21**	.06
Artificial Intelligence can:				
S6	improve healthcare	-.04	-.06	**.10**

Table 2. Persuasion likelihood – the difference in opinions before and after interacting with our visualization. Color scheme: no change (blue), negative change (magenta), positive change (green). Participants changed their opinions on the emotional aspects of conditions (S2), the use of social media data (S4 and S5), and on the use of AI (S6) in healthcare.
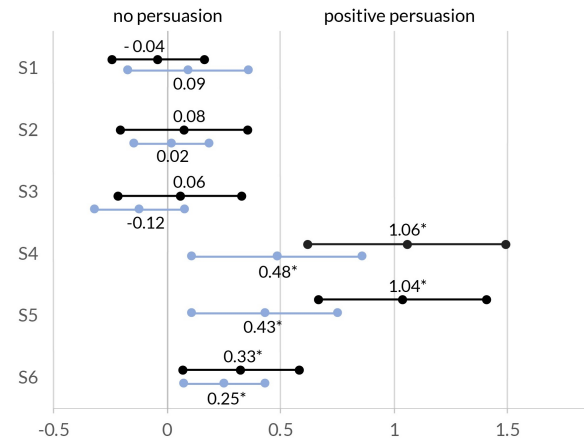

Fig. 8. Statement-wise average opinion change with a 95% confidence. The values for those exposed to our visualization are marked in black, and for those exposed to the baseline in blue. The opinion changes on the three last statements were statistically significant and, as such, are marked with a *.

5.3.3 Answers to the Open-ended Questions (Step 4)

This section summarizes the reasons why the participants decided to change or not their opinions.

To begin with, to test whether our study with domain experts correctly identified the needs of the general public, we asked the baseline participants (those being exposed to the control) how the bio-psycho-social model should be visualized. The answers of both groups turned out to be aligned. Purely based on reading the Wikipedia and Reddit posts, baseline participants mentioned they would like to see a *"list of symptoms belonging to each subreddit, automatically derived from mining posts"* and *"how common symptoms are."* Regarding the psychological side, the baseline participants suggested to show *"the emotions of people in relation to their diagnosis."* Finally, they suggested the social side could be presented as *"a geographical map of the prevalence of diseases."*

By then comparing the baseline participants with the treatment participants, we found three emerging themes.

Bio-psycho-social Model. The most convincing visual elements were the *body part visualization* (4 mentions), the *bubbles of symptoms and diseases* (4 mentions), and the *radar charts* with socio-economic conditions (2 mentions). Similarly, in the baseline group, participants were also influenced by the emotions expressed in the posts (5 mentions), yet, the symptoms (2 mentions), and socio-economic conditions (1 mention) were less important to them.

The main reasons for not changing opinion, on the other hand, were two: the participants either already agreed with the benefits of the bio-psycho-social model (*"I already felt this way and the visualization solidified this."*), or trusted the dominant medical model and, as such,

were reluctant to change (*"Bio-medical model which doctors use is very trustworthy, or at least gives this image."*).

Social Media Data for Healthcare. In relation to social media use, the reasons for adopting a more favourable view included the presence of the *body-part model and emotion ranking* (4 mentions), and of the *bubbles of symptoms and conditions* (2 mentions). One participant was positively impressed by *the variety of social media data* used in the project. The interaction with the psychological layer had a particular effect on our participants (*"it made sense that social media can be considered a record of mental state over time"*, and *"people struggle when put in a position where they have to verbalise their feelings to a doctor, they seem to talk more freely on social media these days."*). The reasons for not changing opinion, on the other hand, were two: the participants were skeptical towards data sources (*"It just makes sense and could be very helpful if the data collected is real."*), or they were concerned about privacy issues (*"Although in my opinion I moved to be less against using social media I still feel the privacy issues are more important."*). The baseline participants mentioned privacy issues more often than the participants exposed to our visualization, likely because the former were exposed to the raw data.

AI for Healthcare. In relation to AI use, the reasons for changing opinions included the visualization elements, such as the *map in combination with the radar chart* (4 mentions), the *added value of merging social media data with official data sources* (3 users), the *bubbles of symptoms and diseases* (2 users), the *body part visualization* (2 users), and the *overall story* flow (1 user). The reasons for not changing opinion, on the other hand, were that the participants had been exposed to the benefits of AI for healthcare before (*"This technology [AI] is already a huge part of the healthcare."*), or they required a proof of reliability of AI in healthcare (*"this part of AI is still to be checked if it works 100%."*).

6 DISCUSSION AND CONCLUSION

We provided a way of visually communicating complex and multi-faceted health information and found that our approach based on a layered narrative was effective not only in creating user awareness but also in promoting opinion change. Our results speak to the importance of visualizations in research fields beyond InfoVis. Take Data Science, for example. This field uses the very same AI methods and systems presented in this paper, yet often neglects information visualization aspects. Yet, these aspects are increasingly becoming critical, from introducing "humans in the loop" in the production of training data, to presenting model predictions in adaptive ways.

6.1 Limitations

Our work comes with two main limitations:

Biases of social media data. While social media data has proven to be a rich and valuable source of health-related information, it carries inherent biases in terms of the non-representative user bases, self-selection biases, and the existence of bots (i.e., non-human users).

Testing a single type of visual storytelling. We have evaluated our Layered Martini Glass visualization against a text-based baseline. However, other types of visual approaches could have been adopted.

6.2 Theoretical implications

We identified two main theoretical aspects that the community could continue to explore in the future:

Adaptive narrative visualization. Health knowledge is distributed differently across different people and roles. Depending on a user's knowledge, a future visualization could automatically adapt the narrative it shows (e.g., a street-level person could be exposed to a different narrative than that of a doctor).

Visual communication with human-in-the-loop. Training data (e.g., social media data) is inherently biased. One way of partly tackling that problem is to introduce "humans in the loop" in the form of, for example, *i)* professionals flagging any incorrect input data or AI results; and *ii)* patients providing training data based on their personal health experiences.

6.3 Practical implications

We identified two main practical implications of our work:

Enhancing healthcare communication. The participants of the study saw our visualization as a new medium of communication between doctors and patients: *"doctors will have a reference in each area to ask about other symptoms and problems."*, and *"as people seem to talk more freely on social media these days"* the verbalisation of feelings to the doctors could *"fast track their diagnosis."*

Visual functionalities for social media platforms. The participants found the visual storytelling superior to the text-based storytelling, which mimicked social-media platform interfaces. This finding suggests that social media platforms could add new functionalities similar to our visualization tool.

REFERENCES

[1] BNF Classifications. https://digital.nhs.uk/data-and-information/areas-of-interest/prescribing/practice-level-prescribing-in-england-a-summary, 2019. [Online; accessed Oct. 2019].

[2] L. M. Aiello, R. Schifanella, D. Quercia, and L. Del Prete. Large-scale and high-resolution analysis of food purchases and health outcomes. *EPJ Data Science*, 8(1):14, 2019.

[3] Y. Albo, J. Lanir, P. Bak, and S. Rafaeli. Off the radar: Comparative evaluation of radial visualization solutions for composite indicators. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):569–578, 2015.

[4] N. Bayat, G. H. Alishiri, A. Salimzadeh, M. Izadi, D. K. Saleh, M. M. Lankarani, and S. Assari. Symptoms of anxiety and depression: A comparison among patients with different chronic conditions. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences*, 16(11):1441, 2011.

[5] A. Bir, N. Freeman, R. Chew, K. Smith, J. Derzon, and T. Day. Making Evidence Actionable: Interactive Dashboards, Bayes, and Health Care Innovation. *eGEMs: The Journal of Electronic Health Data and Methods*, 7(1):40, Aug. 2019.

[6] J. Boy, F. Detienne, and J.-D. Fekete. Storytelling in information visualizations: Does it engage users to explore data? In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1449–1458, Apr. 2015.

[7] J. D. Bradbury and R. E. Guadagno. Documentary narrative visualization: Features and modes of documentary film in narrative visualization. *Information Visualization*, 19(4):339–352, 2020.

[8] K. J. Briant, A. Halter, N. Marchello, M. Escareño, and B. Thompson. The power of digital storytelling as a culturally relevant health promotion tool. *Health promotion practice*, 17(6):793–801, 2016.

[9] C. Bryan, K.-L. Ma, and J. Woodring. Temporal summary images: An approach to narrative visualization via interactive annotation generation and placement. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):511–520, 2016.

[10] J. M. Chambers. *Graphical methods for data analysis*. CRC Press, 2018.

[11] S. Chen, J. Li, G. Andrienko, N. Andrienko, Y. Wang, P. H. Nguyen, and C. Turkay. Supporting story synthesis: Bridging the gap between visual analytics and storytelling. *IEEE Transactions on Visualization and Computer Graphics*, 26(7):2499–2516, July 2020.

[12] M. D. Choudhury and S. De. Mental Health Discourse on Reddit: Self-Disclosure, Social Support, and Anonymity. In *Proceedings of the International AAAI Conference on Weblogs and Social Media*, May 2014.

[13] A. M. Cohen and W. R. Hersh. A survey of current work in biomedical text mining. *Briefings in bioinformatics*, 6(1):57–71, 2005.

[14] D. Concannon, K. Herbst, and E. Manley. Developing a Data Dashboard Framework for Population Health Surveillance: Widening Access to Clinical Trial Findings. *JMIR formative research*, 3(2):e11342, Apr. 2019.

[15] N. S. Consedine and J. T. Moskowitz. The role of discrete emotions in health outcomes: A critical review. *Applied and Preventive Psychology*, 12(2):59–75, 2007.

[16] A. Culotta. Estimating County Health Statistics with Twitter. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1335–1344, Apr. 2014.

[17] H. J. Curtis, R. Croker, A. J. Walker, G. C. Richards, J. Quinlan, and B. Goldacre. Opioid prescribing trends and geographical variation in england, 1998–2018: a retrospective database study. *The Lancet Psychiatry*, 6(2):140–150, 2019.

[18] D. M. Cutler, A. Lleras-Muney, and T. Vogl. Socioeconomic status and health: dimensions and mechanisms. Technical report, National Bureau of Economic Research, 2008.

[19] M. Danesi. *The semiotics of emoji: The rise of visual language in the age of the internet*. Bloomsbury Publishing, 2016.

[20] M. Elias and A. Bezerianos. Exploration Views: Understanding Dashboard Creation and Customization for Visualization Novices. In *Human-Computer Interaction – INTERACT*. Springer Berlin Heidelberg.

[21] G. L. Engel. The need for a new medical model: a challenge for biomedicine. *Science*, 196(4286):129–136, 1977.

[22] A. Faiola, P. Srinivas, and J. Duke. Supporting Clinical Cognition: A Human-Centered Approach to a Novel ICU Information Visualization Dashboard. *AMIA Annual Symposium Proceedings*, 2015:560–569, Nov. 2015.

[23] G. A. Fava and N. Sonino. The biopsychosocial model thirty years later. *Psychotherapy and psychosomatics*, 77(1):1, 2007.

[24] M. Friendly and D. J. Denis. Milestones in the history of thematic cartography, statistical graphics, and data visualization. http://www.datavis.ca/milestones/, 2001. [Online; accessed Oct. 2019].

[25] G. Gonzalez-Hernandez, A. Sarker, K. O'Connor, and G. Savova. Capturing the patient's perspective: a review of advances in natural language processing of health-related text. *Yearbook of Medical Informatics*, 26(01):214–227, 2017.

[26] J. Heer, M. Bostock, and V. Ogievetsky. A tour through the visualization zoo. *Communications of the ACM*, 53(6):59–67, 2010.

[27] J. Hullman, N. Diakopoulos, E. Momeni, and E. Adar. Content, context, and critique: Commenting on a data visualization blog. In *Proceedings of the ACM conference on Computer Supported Cooperative Work & Social Computing*, pp. 1170–1175, Feb. 2015.

[28] H. Kennedy, R. L. Hill, W. Allen, and A. Kirk. Engaging with (big) data visualizations: Factors that affect engagement and resulting new definitions of effectiveness. *First Monday*, 21(11), 2016.

[29] S. S. Khairat, A. Dukkipati, H. A. Lauria, T. Bice, D. Travers, and S. S. Carson. The impact of visualization dashboards on quality of care and clinician satisfaction: integrative literature review. *JMIR human factors*, 5(2):e22, 2018.

[30] J. H. Kim, V. Iyer, S. B. Joshi, D. B. Volkin, and C. R. Middaugh. Improved data visualization techniques for analyzing macromolecule structural changes. *Protein Science*, 21(10):1540–1553, 2012.

[31] J. Lanchantin, R. Singh, B. Wang, and Y. Qi. Deep motif dashboard: visualizing and understanding genomic sequences using deep neural networks. *Pacific Symposium on Biocomputing*, Nov. 2016.

[32] R. Leaman, L. Wojtulewicz, R. Sullivan, A. Skariah, J. Yang, and G. Gonzalez. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. In *Proceedings of the Worshop on Biomedical Natural Language Processing*, pp. 117–125. Association for Computational Linguistics, 2010.

[33] B. Lee, N. H. Riche, P. Isenberg, and S. Carpendale. More than telling a story: Transforming data to visually shared stories. *IEEE Computer Graphics and Applications*, 35(5):84–90, 2015.

[34] H.-C. Lee, Y.-Y. Hsu, and H.-Y. Kao. An enhanced crf-based system for disease name entity recognition and normalization on biocreative v dner task. In *Proceedings of the BioCreative Challenge Evaluation Workshop*, pp. 226–233, 2015.

[35] R. Lin, Z. Ye, H. Wang, and B. Wu. Chronic diseases and health monitoring big data: A survey. *IEEE Reviews in Biomedical Engineering*, 11:275–288, 2018.

[36] S. Liu, D. Maljovec, B. Wang, P. Bremer, and V. Pascucci. Visualizing high-dimensional data: Advances in the past decade. *IEEE Transactions on Visualization and Computer Graphics*, 23(3):1249–1268, 2017.

[37] W.-Y. Liu, B.-W. Wang, J.-X. Yu, F. Li, S.-X. Wang, and W.-X. Hong. Visualization classification method of multi-dimensional data based on radar chart mapping. In *International Conference on Machine Learning and Cybernetics*, vol. 2, pp. 857–862. IEEE, 2008.

[38] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. RoBERTa: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.

[39] N. Ljubešić and D. Fišer. A global analysis of emoji usage. In *Proceedings of the Web as Corpus Workshop*, pp. 82–89, 2016.

[40] K. McCurdy. Visual storytelling in healthcare: Why we should help patients visualize their health. *Information Visualization*, 15(2):173–178, 2016.

[41] S. M. Mohammad and P. D. Turney. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In *Proceedings of the NAACL HLT Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pp. 26–34. Association for Computational Linguistics, 2010.

[42] P. Mylavarapu, A. Yalcin, X. Gregg, and N. Elmqvist. Ranked-list visualization: A graphical perception study. In *Proceedings of the Conference on Human Factors in Computing Systems*, pp. 1–12, 2019.

[43] A. Nikfarjam, A. Sarker, K. O'Connor, R. Ginn, and G. Gonzalez. Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *Journal of the American Medical Informatics Association*, 22(3):671–681, 2015.

[44] L. Nummenmaa, E. Glerean, R. Hari, and J. K. Hietanen. Bodily maps of emotions. *Proceedings of the National Academy of Sciences*, 111(2):646–651, 2014.

[45] F. Ozdemiroglu, C. O. Memis, N. Meydan, B. Dogan, S. M. Kilic, L. Sevincok, and K. Karakus. Self-esteem, pain and suicidal thoughts in a sample of cancer patients. *Psychiatry and Behavioral Sciences*, 7(3):156, 2017.

[46] A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini. The persuasive power of data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2211–2220, 2014.

[47] A. Park and M. Conway. Tracking health related discussions on reddit for public health applications. In *AMIA Annual Symposium Proceedings*, vol. 2017, p. 1362. American Medical Informatics Association, 2017.

[48] A. Park, M. Conway, and A. T. Chen. Examining thematic similarity, difference, and membership in three online mental health communities from reddit: a text mining and visualization approach. *Computers in Human Behavior*, 78:98–112, 2018.

[49] M. J. Paul, A. Sarker, J. S. Brownstein, A. Nikfarjam, M. Scotch, K. L. Smith, and G. Gonzalez. Social media mining for public health monitoring and surveillance. In *Proceedings of the Pacific symposium on Biocomputing*, pp. 468–479. World Scientific, 2016.

[50] D. Perrotta, D. Delle Vedove, C. Obi, R. Pebody, R. Schifanella, and D. Paolotti. Spatio-temporal analysis of flu-related drugs uptake in an online cohort in england. In *Proceedings of the International Conference on Digital Public Health*, pp. 111–118, 2019.

[51] S. Pick, L. H. Goldstein, D. L. Perez, and T. R. Nicholson. Emotional processing in functional neurological disorder: a review, biopsychosocial model and research agenda. *Journal of Neurology, Neurosurgery & Psychiatry*, 90(6):704–711, 2019.

[52] R. Plutchik. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist*, 89(4):344–350, 2001.

[53] G. Ricci, D. Tomassoni, I. Pirillo, A. Sirignano, M. Sciotti, S. Zaami, and I. Grappasonni. Obesity in the european region: social aspects, epidemiology and preventive strategies. *Eur Rev Med Pharmacol Sci*, 22(20):6930–6939, 2018.

[54] A. C. Robinson, U. Demšar, A. B. Moore, A. Buckley, B. Jiang, K. Field, M.-J. Kraak, S. P. Camboim, and C. R. Sluter. Geospatial big data and cartography: research challenges and opportunities for making maps that matter. *International Journal of Cartography*, 3(sup1):32–60, 2017.

[55] P. Robinson, D. Turk, S. Jilka, and M. Cella. Measuring attitudes towards mental health using social media: investigating stigma and trivialisation. *Social psychiatry and psychiatric epidemiology*, 54(1):51–58, 2019.

[56] J. Schwabish. 4 Observations on Animating Your Data Visualizations. https://medium.com/@urban_institute/4-observations-on-animating-your-data-visualizations, 2019. [Online; accessed Aug. 2019].

[57] E. Segel and J. Heer. Narrative visualization: Telling stories with data.

IEEE Transactions on Visualization and Computer Graphics, 16(6):1139–1148, 2010.

[58] J. Suler. The online disinhibition effect. *Cyberpsychology & Behavior*, 7(3):321–326, 2004.

[59] J. Suls and A. Rothman. Evolution of the biopsychosocial model: prospects and challenges for health psychology. *Health psychology*, 23(2):119, 2004.

[60] E. Tutubalina and S. Nikolenko. Combination of deep recurrent neural networks and conditional random fields for extracting adverse drug reactions from user reviews. *Journal of Healthcare Engineering*, 2017, 2017.

[61] L. Van Oudenhove, R. L. Levy, M. D. Crowell, D. A. Drossman, A. D. Halpert, L. Keefer, J. M. Lackner, T. B. Murphy, and B. D. Naliboff. Biopsychosocial aspects of functional gastrointestinal disorders: how central and environmental processes contribute to the development and expression of functional gastrointestinal disorders. *Gastroenterology*, 150(6):1355–1367, 2016.

[62] S. Šćepanović, M.-L. Enrique, D. Quercia, and K. Baykaner. Extracting medical entities from social media. In *ACM Conference on Health, Inference, and Learning*, p. 12. Association for Computing Machinery, 2020.

[63] S. Šćepanović, E. Martín-López, and D. Quercia. Medred dataset. `https://figshare.com/articles/dataset/MedRed/12039609/1`, Mar. 2020.

[64] W. Wang, H. Wang, G. Dai, and H. Wang. Visualization of large hierarchical data by circle packing. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pp. 517–520, 2006.

[65] B. A. Wilbanks and P. A. Langford. A review of dashboards for data analytics in nursing. *CIN: Computers, Informatics, Nursing*, 32(11):545–549, 2014.