


# Towards the Detection and Visual Analysis of COVID-19 Infection Clusters

D. Antweiler<sup>1,2</sup> , D. Sessler<sup>3</sup>, S. Ginzel<sup>1</sup> and J. Kohlhammer<sup>3,4</sup>

<sup>1</sup>Fraunhofer IAIS, Germany <sup>2</sup>Fraunhofer Center for Machine Learning, Germany

<sup>3</sup>Fraunhofer IGD, Germany <sup>4</sup>TU Darmstadt, Germany



**Figure 1:** Visual interface for contact tracing and missing link detection: Our web application consists of two views: On the right the Contact Network shows persons who were in contact with each other (black edges) and persons that we predict to have been in contact (grey, dashed edges). In the center the Epi-Gantt diagram shows events associated with the virus for each person of a selected cluster. At the top and the left the user can apply various filters to select specific groups of persons and control the parameters of the Contact Network.

## Abstract

A major challenge for departments of public health (DPHs) in dealing with the ongoing COVID-19 pandemic is tracing contacts in exponentially growing SARS-CoV2 infection clusters. Prevention of further disease spread requires a comprehensive registration of the connections between individuals and clusters. Due to the high number of infections with unknown origin, the healthcare analysts need to identify connected cases and clusters through accumulated epidemiological knowledge and the metadata of the infections in their database. Here we contribute a visual analytics framework to identify, assess and visualize clusters in COVID-19 contact tracing networks. Additionally, we demonstrate how graph-based machine learning methods can be used to find missing links between infection clusters and thus support the mission to get a comprehensive view on infection events. This work was developed through close collaboration with DPHs in Germany. We argue how our systems supports the identification of clusters by public health experts and discuss ongoing developments and possible extensions.

## CCS Concepts

• **Human-centered computing** → Visual analytics; • **Mathematics of computing** → Graph theory; • **Applied computing** → Health care information systems;

## 1. Introduction

Public health departments (DPHs) are facing great challenges in obtaining and learning from available data on the spread of SARS-CoV2, connecting it to other sources of data and analyzing it regarding ever-pressing questions on how to proceed in combating the virus. Since the outbreak of the novel coronavirus, DPHs have faced a plethora of data that they need to input, compare, consolidate, and recall, yet they lack both the means and the personnel to do so: DPH employees in Germany must search for contact information or compare lists in an established computer system, which sometimes requires additional manual effort, in order to find specific individuals and those with whom they have been in contact and to compare this information. This kind of time-consuming process takes up precious time of both general staff and doctors. Rapid identification and interruption of infection chains through contact tracing plays a key role in responding to the current pandemic and most other infectious diseases e.g. as caused by the norovirus or tuberculosis. To achieve this, DPH staff build and analyze large contact and infection networks, possibly spanning multiple generations of transmissions. By tracing back the source of an infection, individual clusters can be contained and recurring transmission patterns can inform decision making on pandemic responses. However, in practice only around 27% of SARS-CoV-2 cases can be traced back to their source leaving the majority of cases disconnected [BadHA\*20].

This paper reports on a technology that was developed as part of an initiative that started right during the first wave of the SARS-CoV2 pandemic and provided several DPHs in Germany, including Cologne, a region with more than 1 million inhabitants, with visualization and analytics technologies. The federalism in Germany induced a public health setup with more than 370 regional DPHs and the Robert-Koch-Institute (RKI) as a central, federal health institute. We successfully established collaborations, discussed requirements, and continuously presented interim results with 4 DPHs and representatives of RKI during the course of the project. However, our approach is not dependent on a particular public health organization, but is rather based on data that is routinely collected for contact tracing by DPHs and partly shared with the federal RKI via their dedicated system called SurvNet@RKI [KAF\*07]. In this sense, our approach supports cluster detection that is independent of further technical means (e.g. tracing apps) and without further collecting data beyond the information contained in the DPH databases. We first describe a general scenario that crystallized during the collaboration with experts from Germany's largest DPH in Cologne. This will motivate our contributions presented in this paper. We then describe related work in the areas of epidemiological visual analytics and link prediction on complex networks. The proposed visual analytics system is then presented and evaluated with an exemplary inspection.

## 2. Scenario

Motivated by our collaboration with several DPHs in Germany between the two first waves of infections, this section describes a short scenario for the usage of our technique, loosely following the design triangle by Miksch and Aigner [MA14]. Our targeted user group are healthcare analysts at a DPH who regularly trace in-

fections in a broader view and time frame, and need an overview of the spread of infections to inform local authorities and help advise the implementation or adaptation of restrictions and quarantine measures. All DPHs observe and trace single infected citizens, but also look at the broader developments and, especially and most challenging, clusters of infections. Due to the low number of infections that can be traced back to the host of infection, the healthcare analysts try to identify clusters through accumulated epidemiological knowledge and the metadata of the infections in their database. This includes the time of infection, the address of the citizen, and the provided contacts that the citizen can remember during the telephone interview. Learning more about the clusters of infection and understanding the infection routes both promise an important insight into the epidemiological dynamic, which again drive intervention policies and the possibility to strategically prioritize contact tracing and impose (or lift) restrictions. In many DPH regions the number of infections in care homes were particularly high. Finding out whether infections in such facilities were driven by internal infection chains or were continuously induced from outside is very difficult without known infection origins. Therefore, our motivation was to visually provide all the available information in an interactive epidemiological network visualization and a temporal visualization of all contacts within a cluster. In addition, we employed a link prediction algorithm to suggest connections between formerly separated clusters if such a connection can be justified by the data, i.e. the temporal and geographical relation of infected contacts.

## 3. Related work

Our work is closely related to the research areas of visual analytics for public health and infectious diseases [CAD\*14] and link prediction in complex networks.

### 3.1. Visual analytics for public health

There have been approaches in the past that especially detect and visualize infection clusters. Si et al. examined the data of the SARS outbreaks from 2003 onward and developed a prototype system that tracks the patients' histories in detail to find contacts and visiting activities [SLBAF10]. This is a good example for approaches that heavily depend on a fine-grained record about the movement and activities of each infected person and their contacts, which our approach could not rely on. There have been a number of COVID-19 and SARS-CoV2 related publications in the visualization community. For example, Leite et al. use temporal visualization to understand the prediction of different COVID-19 scenarios, but on a higher level of abstraction and not on the contact level [LSC\*20]. Similarly, Afzal et al. focus on geographical visualization to help public health officials prepare and exercise response plans in pandemic outbreak scenarios [AGJ\*20]. Other approaches take a more general epidemiological approach and address emergent pandemics [PBK\*19], the role of social networks in epidemiology [CF09], or rather epidemic management [YDH\*17, AKMR16]. Epidemic management and immediate outbreak control are also the focus of SORMAS [FDA\*15] (originally developed for Ebola outbreaks), which is used in some German DPHs in addition to SurvNet@RKI, but not for the interactive detection of infection clusters with link predictions.

There are a wider range of possibilities for tracking patients and infections as well as finding clusters of infections in hospitals and similar confined environments. The project HiGHmed focused one of its use cases on infection control in hospitals to elicit transmission path of multi-resistant bacteria within and between clinics [MPW\*20], [BPW\*21]. Their literature analysis confirms that there is only little related work on the visual analysis of disease spread in dynamic networks. While the HiGHmed pathway interface also combines network and temporal visualizations, it does not enable the detection of new infection clusters. Instead, our network visualization shows all infection sub-networks, including predicted links, and a temporal visualization that is focused on the characteristics of the COVID-19 viral transmission and quarantine measures.

### 3.2. Link prediction

Link prediction is a common task when dealing with graph-structured data. Any typical dataset may exhibit missing or unobserved edges. Given two nodes in a network, that are not connected through an edge, we want to estimate the probability of this edge being absent based on the currently observed links between nodes and their properties. Regular applications include predicting relationships in social networks [LNK03], classifying unseen triplets in knowledge graphs [WMWG17] or identifying user-product interactions in recommendation systems [vdBKW17]. An additional major area of application is the prediction of links that may or may not appear in dynamic and temporal networks, that change over time [DM20].

Simple heuristics to estimate the probability of a missing link use common neighbors and exploit the topological structure of the graph to assess whether the two nodes belong to the same community and should therefore be connected. Other methods include random walks, factorization of the adjacency matrix of the graph and probabilistic graphical models. In recent years, deep learning based methods such as DeepWalk [PARS14], node2vec [GL16] and Graph Convolutional Networks (GCNs) [KW16] introduced the concept of node embeddings which, together with various similarity metrics, can be used to estimate the likelihood of an edge.

However, most of the discussed approaches are ill-suited for our problem at hand. Firstly, they assume the graph is composed of large connected components and predict edges within these components with far greater accuracy than between them. Contact tracing data in epidemiology takes the form of many smaller components and users are especially interested in identifying connections between disconnected clusters [AIT\*07]. Secondly, the considered models do not account for interpretability of results and are therefore unsuitable for applications in a risk-averse healthcare setting.

Overall, to the best of our knowledge, there are no references that combine link prediction in infection networks with time-oriented visualization of epidemiological contacts. Therefore, in the following section we propose a simple, interpretable, and accurate approach to identify possible missing edges in a given contact tracing network.

## 4. Proposed Visual Analytics System

Our visual analytics system consists of a machine learning module that predicts missing links in the contact data and a web application

that visualizes this data enriched with the prediction results. We use the anonymized data from Cologne's health department as the basis for both, the machine learning and the visualization system. In the following, we present the two main components of our visual analytics system in detail.

### 4.1. Predicting missing links

We predict missing edges with high probability inside our contact tracing network based on the known connections and their attributes. Feature selection was based on DPH experts' feedback to find feasible real-world features that are most relevant for our task. Given an ordered pair of nodes  $(u, v)$  inside our graph, we calculate the *days between infection reports*, the *difference in years of age* and a binary feature that encodes if  $u$  and  $v$  live in the same *housing block*, an anonymized proxy for the home address. The generated feature vectors are the input to multiple different classification models including *DecisionTrees*, *RandomForest*, *NaiveBayes* and *Support Vector Classification*. Hyperparameters are selected by performing a grid search and cross-validation over a predefined parameter space.

We evaluate our approach on the anonymized data from Cologne's containing 44.634 persons with 11.652 'index cases' who were positively tested for COVID-19 among them. For 7.846 index cases (67.3%) the infection source is unknown or not part of the dataset. To create a data set for training our models, we select all observed edges and add 40,000 negative samples of random pairs of persons, who do not share a connection inside our dataset. We perform a random 85%/15% train-test split of the resulting dataset and train our model to predict the probability of an infection. For model performance comparison, we compute the receiver operating characteristic (ROC) curve for all classifiers plotting true positive rate (TPR) versus false positive rate (FPR) on the test set. Results are reported in Table 1. Area under the ROC curve (AUROC) is best for the RandomForest model with 0.98, but all other models do achieve similar results. For the presented use case, a high precision is particularly desirable as resources to track additional infection transmissions at the DPH might be scarce and false positive predictions could lead to negative consequences (e.g. quarantine) for the citizens involved.

We investigate additional features but note that no statistically significant gains in performance are achieved by including them. To retain high interpretability of results, we confine the model to the attributes selected above and integrate the best performing DecisionTree model into our visualization system.

### 4.2. Visualization system

Our visual interface consists of two main interactive views. In order to facilitate a fast identification of infection clusters we implemented a graph-based view of the contacts in the dataset. Additionally, we combine this graph view with an event-based visualization that helps to analyze the temporal characteristics of the infection dynamic. Both visualizations are linked and allow to select groups of persons for a more detailed analysis. Additionally, the visualization system provides multiple filtering capabilities to support the analysis tasks. We use a JavaScript web frontend implemented with

the framework react.js [Rea21b] and material-ui [Mat21]. For the graph visualization we use the library react-force-graph [Rea21a] that implements a force directed layout. The event based visualization is implemented with d3.js [BOH11].

At the right of Figure 1 the graph based visualization shows the contact network. Each node represents a person and describes their state, which is color-coded according to the collected data. This color-coding allows for a fast indication whether a person was tested positive or had symptoms prior or shortly after the contact with other people. The directed edges show contacts between people and thus indicate potential infection paths of the virus. The solid, black edges represent the registered contacts from the data set. Additional dashed, grey edges represent the predicted links according to their calculated probability, ranging from 100% (solid grey line) to 50% (widely dashed line). We decided to not show predicted links below 50% probability. The main benefit of the node link diagram is that it allows to identify individuals that were in contact with each other at a glance. Therefore, it speeds up the identification of infection chains which is crucial for timely countermeasures. Further, the predicted links make it possible to identify additional potential contacts, which can be verified by the health department staff to fill the missing links in their data. Thus, a fast and more profound assessment can be made about the extent of infection clusters, which should lead to more informed decisions to contain the spread of the virus.

The *Epi-Gantt* diagram in the center of Figure 1 shows event records of connected persons. Each row in this view represents a person and lists important time-dependent events, like contacts with other people, start of symptoms or the outcome of a SARS-CoV-2 test. Bars represent continuous events, and circles represent events that can be considered a point in time. Detailed information about these events can be displayed via a tooltip by hovering the mouse over the corresponding bar or circle. This visualization can help to get an insight in the sequence of events that frequently lead to the propagation of the infection. These insights about the typical progress of the infection in a specific region might help to pinpoint weaknesses in the containment strategy and thus should lead to better countermeasures in the future. The *Epi-Gantt* diagram can also be used to show clusters of persons connected by predicted links, which can be used to compare the event sequences of confirmed infection paths with the predicted ones. This way, the plausibility of the proposed missing links can be further investigated in the temporal visualization.

To facilitate an efficient contact tracing and analysis of the data we implemented various filters. At the top of Figure 1 is the time filter which can be used to narrow down the displayed data to the time frame of interest. On the left are additional filters that can be used to specify a person (via a pseudo ID) or to manipulate the minimum sizes of displayed contact networks. Further, it is possible to display only networks and events of persons that had a positive SARS-CoV-2 test result or had symptoms associated with the virus. It is also possible to adjust the amount of displayed predicted links by adjusting the minimum probability threshold via a slider control.

Model architecture	DT	RF	NB	SVC
AUROC	0.97	<b>0.98</b>	0.96	0.94

**Table 1:** Classification performance of trained models on the contact tracing benchmark dataset. Reported is the area under receiver operating characteristic (AUROC, higher is better).

#### 4.3. Discussion

We discussed our proposed visual analytics system and its features with healthcare experts from multiple DPHs to evaluate the benefits for the established use cases. The feedback was overwhelmingly positive. A combined view of the temporal epidemiological data via the time scale together with the network-based view of the disease spreading between infected people was perceived as supportive to comprehend the spatio-temporal dynamic of the pandemic. Additionally, the presentation of possible missing edges was perceived as helpful in the detection of plausible connections between known clusters. Identifying infection sources and tracing the spread of a virus through communities is especially important with new mutations of SARS-CoV-2 arising. However, the current system is limited to the contact data from individual DPHs, while in practice infection clusters can span across multiple jurisdiction that cover different populations, cities and countries. This limitation could be addressed with the availability of standardized contact tracing data and would require sophisticated data and visualization architecture that allows the rendering of possibly millions of nodes. Additionally, the current link prediction model is driven by routinely collected observations. While this allows a flexible use of our system, it is important to note that modelling the effect of changes in epidemiological parameters as well as public restrictions (e.g. in care homes) over time may provide further, more detailed insights.

#### 5. Conclusion and future work

We presented a novel visual analytics approach to detect SARS-CoV2 infection clusters. Our system calculates and visualizes possible missing infection routes inside the contact tracing network and supports the analysis of time-dependent events that led to the spread of the virus. The evaluation together with local healthcare authority experts shows the benefits of our approach. In future work we would like to extend our framework with model explanations to help better understand the complex interactions underlying SARS-CoV2 infection clusters. Based on the presented results, link prediction methods together with larger and more detailed contact tracing data, while still preserving privacy, could potentially achieve even better visual and analytical performance and support the control of pandemics.

#### 6. Acknowledgments

We would like to thank the DPH experts for valuable discussions and the insights into their work that motivated our system. This work has been partly funded by the Fraunhofer Cluster for Cognitive Internet Technologies, the BMBF Competence Center for Machine Learning ML2R (01IS18038B), and the Fraunhofer vs. Corona initiative.



## References

- [AGJ\*20] AFZAL S., GHANI S., JENKINS-SMITH H. C., EBERT D. S., HADWIGER M., HOTEIT I.: A visual analytics based decision making environment for covid-19 modeling and visualization. In *2020 IEEE Visualization Conference (VIS)* (2020), pp. 86–90. doi:10.1109/VIS47514.2020.00024. 2
- [AIT\*07] ANDRE M., IJAZ K., TILLINGHAST J., KREBS V., DIEM L., METCHOCK B., CRISP T., MCELROY P.: Transmission network analysis to complement routine tuberculosis contact investigations. *American Journal of Public Health* 97, 3 (2007), 470–477. cited By 38. doi:10.2105/AJPH.2005.071936. 3
- [AKMR16] ABDELHAMID S. E., KUHLMAN C. J., MARATHE M. V., RAVI S. S.: Interactive exploration and understanding of contagion dynamics in networked populations. In *2016 International Conference on Behavioral, Economic and Socio-cultural Computing (BESC)* (2016), pp. 1–6. doi:10.1109/BESC.2016.7804480. 2
- [BadHA\*20] BUDA S., AN DER HEIDEN M., ALTMANN D., DIERCKE M., HAMOUDA O., REXROTH U.: Infektionsumfeld von erfassten covid-19-ausbrüchen in deutschland. *RKI Epidemiologisches Bulletin*, 38 (2020), 3–12. doi:http://dx.doi.org/10.25646/7093. 2
- [BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D<sup>3</sup> data-driven documents. *IEEE Transactions on Visualization & Computer Graphics*, 12 (2011), 2301–2309. URL: <https://d3js.org/>. 4
- [BPW\*21] BAUMGARTL T., PETZOLD M., WUNDERLICH M., HOHN M., ARCHAMBAULT D., LIESER M., DALPKE A., SCHEITHAUER S., MARSCHOLLEK M., EICHEL V. M., MUTTERS N. T., CONSORTIUM H., LANDESBERGER T. V.: In Search of Patient Zero: Visual Analytics of Pathogen Transmission Pathways in Hospitals. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (Feb. 2021), 711–721. URL: <https://ieeexplore.ieee.org/document/9286903/>, doi:10.1109/TVCG.2020.3030437. 3
- [CAD\*14] CARROLL L. N., AU A. P., DETWILER L. T., CHIEH FU T., PAINTER I. S., ABERNETHY N. F.: Visualization and analytics tools for infectious disease epidemiology: A systematic review. *Journal of biomedical informatics* 51 (2014), 287–298. doi:<https://doi.org/10.1016/j.jbi.2014.04.006>. 2
- [CF09] CHRISTAKIS N. A., FOWLER J. H.: Social Network Visualization in Epidemiology. *Norsk epidemiologi = Norwegian journal of epidemiology* 19, 1 (2009), 5–16. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3337680/>. 2
- [DM20] DIVAKARAN A., MOHAN A.: Temporal Link Prediction: A Survey. *New Generation Computing* 38, 1 (Mar. 2020), 213–258. URL: <https://doi.org/10.1007/s00354-019-00065-z>, doi:10.1007/s00354-019-00065-z. 3
- [FDA\*15] FÄHNRICH C., DENECKE K., ADEOYE O. O., BENZLER J., CLAUS H., KIRCHNER G., MALL S., RICHTER R., SCHAPRANOW M. P., SCHWARZ N., TOM-ABA D., UFLACKER M., POGGENSEE G., KRAUSE G.: Surveillance and Outbreak Response Management System (SORMAS) to support the control of the Ebola virus disease outbreak in West Africa. *Eurosurveillance* 20, 12 (Mar. 2015). URL: <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES2015.20.12.21071>, doi:10.2807/1560-7917.ES2015.20.12.21071. 2
- [GL16] GROVER A., LESKOVEC J.: node2vec: Scalable Feature Learning for Networks. *arXiv:1607.00653 [cs, stat]* (July 2016). arXiv: 1607.00653. URL: <http://arxiv.org/abs/1607.00653>. 3
- [KAF\*07] KRAUSE G., ALTMANN D., FAENSEN D., PORTEN K., BENZLER J., PFOCH T., AMMON A., KRAMER M. H., CLAUS H.: SurvNet Electronic Surveillance System for Infectious Disease Outbreaks, Germany. *Emerging Infectious Diseases* 13, 10 (Oct. 2007), 1548–1555. URL: [http://wwwnc.cdc.gov/eid/article/13/10/07-0253\\_article.htm](http://wwwnc.cdc.gov/eid/article/13/10/07-0253_article.htm), doi:10.3201/eid1310.070253. 2
- [KW16] KIPF T. N., WELLING M.: Semi-supervised classification with graph convolutional networks. *arXiv:1609.02907* (2016). 3
- [LNK03] LIBEN-NOWELL D., KLEINBERG J.: The link prediction problem for social networks. In *Proceedings of the twelfth international conference on Information and knowledge management* (New Orleans, LA, USA, Nov. 2003), CIKM '03, Association for Computing Machinery, pp. 556–559. URL: <https://doi.org/10.1145/956863.956972>, doi:10.1145/956863.956972. 3
- [LSC\*20] LEITE R. A., SCHETINGER V., CENEDA D., HENZ B., MIKSCH S.: Covis: Supporting temporal visual analysis of covid-19 events usable in data-driven journalism. In *2020 IEEE Visualization Conference (VIS)* (2020), pp. 56–60. doi:10.1109/VIS47514.2020.00018. 2
- [MA14] MIKSCH S., AIGNER W.: A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics* 38 (2014), 286–290. 2
- [Mat21] MATERIAL UI: React components that implement Google's Material Design, 2021. URL: <https://material-ui.com/>. 4
- [MPW\*20] MÜLLER M., PETZOLD M., WUNDERLICH M., BAUMGARTL T., HÖHN M., EICHEL V., MUTTERS N. T., SCHEITHAUER S., MARSCHOLLEK M., LANDESBERGER T. V.: Visual Analysis for Hospital Infection Control using a RNN Model. In *EuroVis Workshop on Visual Analytics (EuroVA)* (2020), Turkay C., Vrotsou K., (Eds.), The Eurographics Association. doi:10.2312/eurova.20201090. 3
- [PARS14] PEROZZI B., AL-RFOU R., SKIENA S.: DeepWalk: Online Learning of Social Representations. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining* (Aug. 2014), 701–710. arXiv: 1403.6652. URL: <http://arxiv.org/abs/1403.6652>, doi:10.1145/2623330.2623732. 3
- [PBK\*19] POLONSKY J. A., BAIDJOE A., KAMVAR Z. N., CORI A., DURSKE K., EDMUNDS W. J., EGGO R. M., FUNK S., KAISER L., KEATING P., DE WAROUX O. L. P., MARKS M., MORAGA P., MORGAN O., NOUVELLET P., RATNAYAKE R., ROBERTS C. H., WHITWORTH J., JOMBART T.: Outbreak analytics: a developing data science for informing the response to emerging pathogens. *Philosophical Transactions of the Royal Society B: Biological Sciences* 374, 1776 (July 2019), 20180276. doi:10.1098/rstb.2018.0276. 2
- [Rea21a] REACT-FORCE-GRAPH - VASTURIANO, 2021. URL: <https://github.com/vasturiano/react-force-graph>. 4
- [Rea21b] REACT.JS - FACEBOOK INC., 2021. URL: <https://reactjs.org/>. 4
- [SLBAF10] SI Y.-W., LEONG K.-I., BIUK-AGHAI R. P., FONG S.: Hidden cluster detection for infectious disease control and quarantine management. In *Visual Information Communication* (Boston, MA, 2010), Huang M. L., Nguyen Q. V., Zhang K., (Eds.), Springer US, pp. 159–172. 2
- [vdBKW17] VAN DEN BERG R., KIPF T. N., WELLING M.: Graph convolutional matrix completion, 2017. *arXiv:1706.02263*. 3
- [WMWG17] WANG Q., MAO Z., WANG B., GUO L.: Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering* 29, 12 (2017), 2724–2743. doi:10.1109/TKDE.2017.2754499. 3
- [YDH\*17] YAÑEZ A., DUGGAN J., HAYES C., JILANI M., CONNOLLY M.: Pandemcap: Decision support tool for epidemic management. In *2017 IEEE Workshop on Visual Analytics in Healthcare (VAHC)* (2017), pp. 24–30. doi:10.1109/VAHC.2017.8387497. 2