

## Original Article

# Transmission visualizations of healthcare infection clusters: A scoping review

Mya B. Brady MPH<sup>1,2</sup> , Helena M. VonVille MLS MPH<sup>3</sup> , Joseph F. White MPH<sup>4</sup>, Elise M. Martin MD, MS<sup>2,5</sup> ,  
Nathan J. Raabe MPH<sup>2,4</sup> , Julie M. Slaughter MBA<sup>1</sup> and Graham M. Snyder MD, SM<sup>1,2</sup> 

<sup>1</sup>Department of Infection Prevention and Control, UPMC Presbyterian–Shadyside, Pittsburgh, Pennsylvania, <sup>2</sup>Division of Infectious Diseases, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, <sup>3</sup>University of Pittsburgh Health Sciences Library System, Pittsburgh, Pennsylvania, <sup>4</sup>Department of Epidemiology, University of Pittsburgh School of Public Health, Pittsburgh, Pennsylvania and <sup>5</sup>Veterans' Affairs Pittsburgh Healthcare System, Pittsburgh, Pennsylvania

### Abstract

**Objective:** To evaluate infectious pathogen transmission data visualizations in outbreak publications.

**Design:** Scoping review.

**Methods:** Medline was searched for outbreak investigations of infectious diseases within healthcare facilities that included  $\geq 1$  data visualization of transmission using data observable by an infection preventionist showing temporal and/or spatial relationships. Abstracted data included the nature of the cluster(s) (pathogen, scope of transmission, and individuals involved) and data visualization characteristics including visualization type, transmission elements, and software.

**Results:** From 1,957 articles retrieved, we analyzed 30 articles including 37 data visualizations. The median cluster size was 20.5 individuals (range, 7–1,963) and lasted a median of 214 days (range, 12–5,204). Among the data visualization types, 10 (27%) were floor-plan transmission maps, 6 (16%) were timelines, 11 (30%) were transmission networks, 3 (8%) were Gantt charts, 4 (11%) were cluster map, and 4 (11%) were other types. In addition, 26 data visualizations (70%) contained spatial elements, 26 (70%) included person type, and 19 (51%) contained time elements. None of the data visualizations contained contagious periods and only 2 (5%) contained symptom-onset date.

**Conclusions:** The data visualizations of healthcare-associated infectious disease outbreaks in the systematic review were diverse in type and visualization elements, though no data visualization contained all elements important to deriving hypotheses about transmission pathways. These findings aid in understanding the visualizing transmission pathways by describing essential elements of the data visualization and will inform the creation of a standardized mapping tool to aid in earlier initiation of interventions to prevent transmission.

(Received 22 March 2022; accepted 10 May 2022)

To stop pathogen transmission during a healthcare-associated outbreak, transmission pathways need to be hypothesized.<sup>1</sup> Once potential routes of transmission are defined, whether they are probable or confirmed through person, place, and time, intervention measures can be initiated or continued to interrupt pathogen spread. Transmission pathways may be complex; the type of contact that occurred and the time and space in which the event occurred may not be completely clear. Compared to traditional epidemiological case investigations using only line-list data, visualizing

transmission pathways can improve hypothesized relationships in time and space and can, therefore, promote more timely initiation of infection prevention and control measures.<sup>2</sup>

No standard data visualization tool exists for describing transmission pathways. Typically, line lists comprising patient characteristics, event dates, and elements of clinical care provided are used to generate hypotheses of transmission pathways, with or without accompanying bed traces. Infection surveillance software allows users to review admission locations during the hospital stay, which can provide basic information about shared time and space and therefore transmission opportunities. However, transmission may occur even when patients do not share a proximate admission location at, or close to, the same time.<sup>3</sup>

Tools for data visualization of transmission pathways have been described for use by healthcare-associated infection prevention teams for cluster investigation, and the extent of use these tools is not known. Two examples include Dotmapper and HospMapper: Dotmapper was developed by researchers as an

**Author for correspondence:** Graham M. Snyder, MD, MS, Falk Medical Building, 3601 Fifth Avenue, Suite 150 Pittsburgh, PA 15213. E-mail: [snydergm3@upmc.edu](mailto:snydergm3@upmc.edu)

**PREVIOUS PRESENTATION.** Portions of the data in this manuscript were presented and published through the University of Pittsburgh School of Public Health on the Institutional Repository at the University of Pittsburgh (D-Scholarship) and embargoed until January 6, 2024. The protocol was published on Open Science Framework (<https://osf.io/azuyr/>).

**Cite this article:** Brady MB, et al. (2022). Transmission visualizations of healthcare infection clusters: A scoping review. *Antimicrobial Stewardship & Healthcare Epidemiology*, <https://doi.org/10.1017/ash.2022.237>

© The Author(s), 2022. Published by Cambridge University Press on behalf of The Society for Healthcare Epidemiology of America. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (<http://creativecommons.org/licenses/by-nc-sa/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is used to distribute the re-used or adapted article and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use.

easily accessible interactive mapping tool to visualize infection clusters over large geographic regions using Geographic Information Systems (GIS) (Esri, California).<sup>4</sup> HospMapper was developed for nosocomial transmission and can show data for patients who have a positive test for a pathogen, with outputs of an epidemic curve, a timeline, and a schematic ward plan.<sup>5</sup>

Furthermore, no guidelines or standards have been established to evaluate data visualization tools for transmission events. Salinas *et al*<sup>6</sup> provided advice for best practices when creating a data visualization. In a systematic review published in 2014, Carroll *et al*<sup>7</sup> found that usability and utility were important functions when determining helpfulness of a data visualization tool. Davis *et al*<sup>8</sup> contemporaneously reported that 15% of published outbreak investigations included spatiotemporal analyses.

In infection prevention and hospital epidemiology, there is a need not only to describe the data visualizations that can be used to generate hypotheses about transmission pathways but also a critical analysis of the characteristics of each data visualization, as well as the software used to create them. Such an analysis would allow infection prevention and hospital epidemiology teams to choose a most effective data visualization tool for the cluster being investigated, to generate hypotheses about transmission pathways more effectively and therefore improve cluster investigations, to implement interventions to interrupt transmission more quickly to ultimately better prevent patient harm.

In this scoping review, we have evaluated the infectious-pathogen transmission data visualizations used in healthcare settings to determine which methods and elements of visualizing transmission in healthcare settings are effective at generating hypotheses regarding transmission pathways.

## Methods

### *Published evidence eligibility criteria*

Articles eligible for inclusion in the scoping review were peer reviewed, written in English, and published after 1985 (when infection preventionists (IPs) first documented using electronic surveillance including electronic medical records and information technology tools that allow for automated data collection).<sup>9,10</sup> All eligibility criteria were selected in the search algorithm.

Articles were excluded if they did not contain patient data with a methodology appropriate for evaluating a healthcare-associated infectious disease or infectious pathogen cluster, such as observational cluster investigation methods.<sup>11</sup> Therefore, cost-benefit analyses, meta-analyses, randomized control trials, community outbreaks, and other nonhealthcare studies were not included in the review.

Full article review used the following 4 inclusion criteria to select screened articles for data abstraction: (1) Infectious diseases met healthcare-associated infection definitions using National Healthcare Safety Network criteria (NHSN)<sup>12</sup> but could be an infectious disease not defined using NHSN healthcare-associated infection criteria if  $\geq 1$  transmission event occurred within a healthcare facility. This was done to capture all potential pathogen transmission in the healthcare setting. (2) The source described an infectious disease cluster. Multiple clusters in an outbreak or publication were included if each cluster was contained in 1 healthcare facility. (3) The results contained 1 or more data visualizations demonstrating pathogen transmission pathways with data observable by an infection preventionist showing temporal and/or spatial relationships using patient health data and epidemiological data.

For example, data visualizations could not contain only a phylogenetic tree, which alone would not include various person type data (ie, patients, healthcare workers) or test positive dates; however, a phylogenetic tree could serve as a data visualization if it incorporated other nongenetic epidemiologically relevant data. (4) The facilities where transmission occurred was considered a healthcare facility. Healthcare facilities were defined as structures whose primary purpose was designed to provide care to individuals where person-to-person transmission can occur and where healthcare personnel are present. Healthcare facilities include long-term care facilities, acute-care facilities, rehabilitation facilities, and behavioral health facilities.

### *Information sources and search strategy*

Medline (Ovid) was searched by a health sciences librarian (H.M.V.V.) with systematic review experience. The date of the search was August 24, 2021. Concepts that comprised the search were healthcare-acquired infections and cluster analysis or geographic mapping. A combination of MeSH terms and title, abstract, and keywords was used to develop the Medline search that was checked against a known set of studies (Supplementary Table S1 online). Articles were also considered for review if they were cited as references within the papers found through the original Medline search and appeared to have relevance for this review.

Citations were uploaded from EndNote (Clarivate, UK) to DistillerSR (Evidence Partners, Canada) for the study selection process. All study selection decisions were stored in DistillerSR. Two authors (M.B.B. and J.F.W.) conducted the title, abstract, and full-article eligibility screening. When a consensus on article eligibility could not be reached, a third investigator (G.M.S.) independently reviewed the citation, abstract, and manuscript to determine eligibility. The data charting was completed by 3 reviewers (M.B.B., J.F.W., and N.J.R.) for articles that passed the full-text screening process. The protocol was published on Open Science Framework prior to beginning the literature search.<sup>13</sup>

### *Data abstraction*

Data abstraction was completed independently by 3 reviewers (M.B.B., J.F.W., and N.J.R.) from the eligible articles using DistillerSR using a data abstraction form developed by the investigators. The data abstracted included the nature of the cluster(s) including pathogen, scope of transmission, and individuals involved. Characteristics of the transmission pathway data visualization(s) included type of visualization, elements of transmission incorporated into the visualization, and software used in the visualization (Supplementary Table S2).

### *Data visualization categorization*

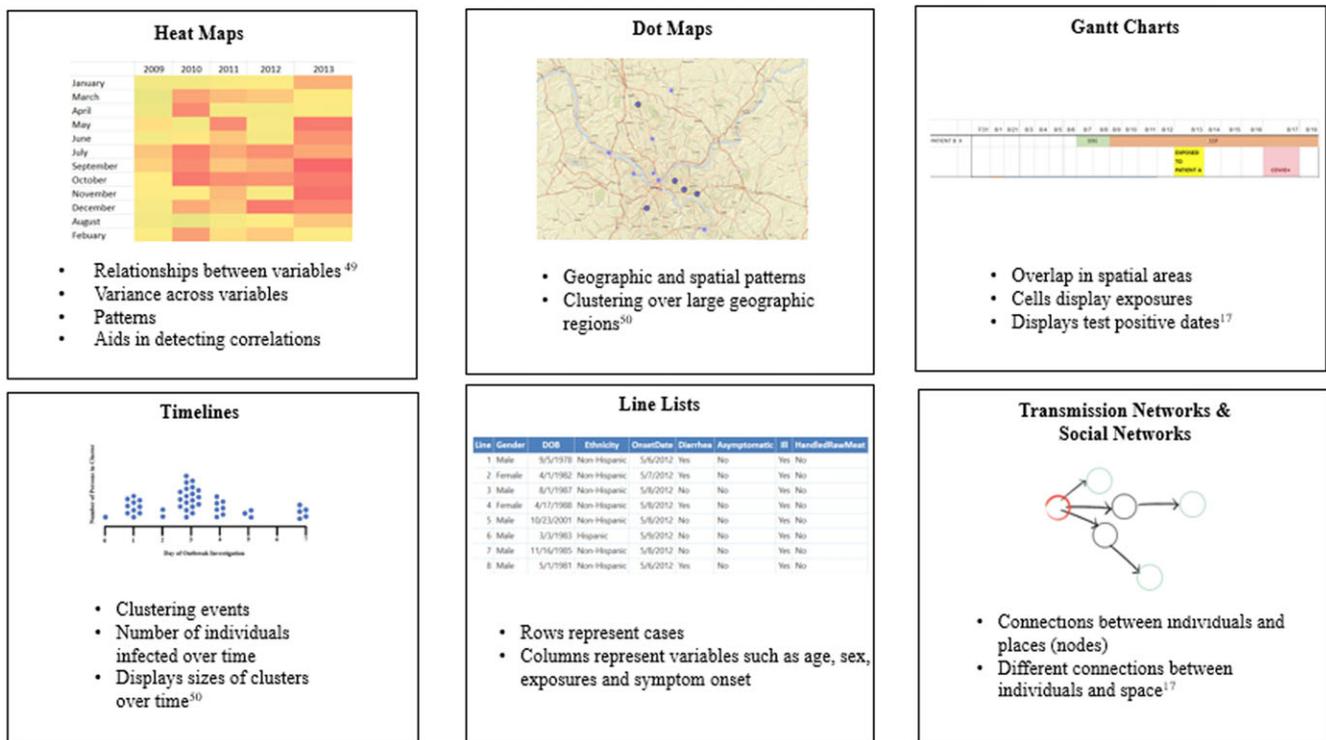
Data visualization categories were defined data visualization types within specified categories of hypothetical examples (Fig. 1). If data visualizations were unable to be grouped within the established example types, new categories were added to ensure that all data visualizations were incorporated and analyzed.

### *Statistical analysis*

Statistical analysis, the calculations of means, medians, and totals, were performed using Excel. Characteristics of the studies and data visualizations are described using aggregate averages.

The project, including a subsequent evaluation and implementation of a software tool for data visualization, underwent formal

## Types of Data Visualizations



**Fig. 1.** Examples of data visualization types from reported healthcare-associated infectious diseases outbreaks. Note: Citations for information for the data visualizations: heat map,<sup>49</sup> dot maps,<sup>50</sup> timelines,<sup>50</sup> Gantt chart, transmission networks, and social networks.<sup>17</sup> All visualizations are hypothetical examples created by the author this year.

review and was granted ethical approval (Project 3405) as a quality improvement project by University of Pittsburgh Medical Center's Quality Improvement Review Committee.

## Results

### Study characteristics

With the exclusion of 1 duplicate article, 1,957 articles were identified through the Ovid Medline database search (Fig. 2) and 1,707 (87.2%) articles were included after review of the citation and abstract. Among the articles that were excluded, 335 did not include an infectious pathogen, 829 were not human studies, and 1,691 used nonpertinent study methodology. The remaining 250 (12.7%) articles underwent full text review, of which 220 did not meet  $\geq 1$  inclusion criteria: 13 did not study an NHSN-defined or healthcare-associated organism, 95 did not describe an infectious diseases cluster, 200 did not include  $\geq 1$  data visualization, 40 did not occur in a healthcare facility, and 1 reported an outbreak in  $>1$  facility. The remaining 30 articles (1.5%) fulfilled the eligibility criteria and were included in data abstraction and analysis.

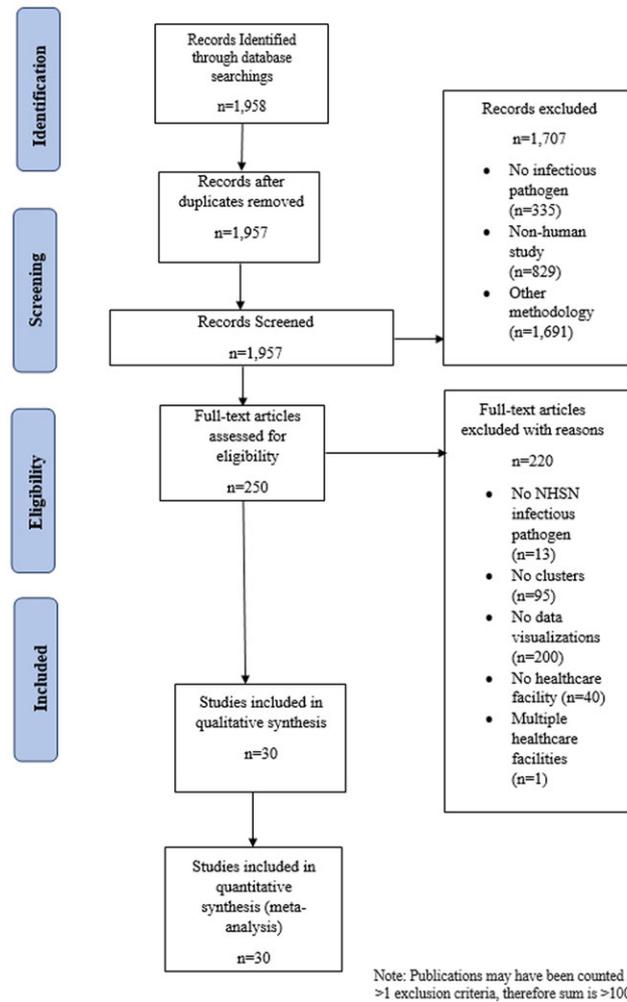
The 30 analyzed studies were conducted in healthcare facilities in North America,<sup>14–18</sup> South America,<sup>19</sup> Europe,<sup>20–31</sup> Africa,<sup>32,33</sup> Asia,<sup>34–42</sup> and Australia<sup>43</sup> (Table 1). Fourteen studies (46.7%) described outbreaks occurring in acute-care facilities<sup>15,16,19,21,25,29–32,34–38</sup> and 11 (36.7%) studies were conducted in a tertiary-care hospitals.<sup>14,18,22–24,27,28,33,40,42,43</sup> Of the remaining studies, 2 studies took place in a community hospital,<sup>17,41</sup> 2 studies took place in a geriatric hospital,<sup>26,39</sup> and 1 study took place in a

long-term care facility with a rehabilitation care ward.<sup>20</sup> Of the 30 studies, 17 (56.6%) were prospective studies.<sup>15,17,21,23,24,26,28–31,33–35,38,40–42</sup>

The 30 studies described clusters with a median cluster size of 20.5 individuals (range, 7 to 1,963) and a median duration of 214 days (range, 12–5,204) (Table 1). Also, 15 (50%) of the causative pathogens described in clusters were viral,<sup>14,15,17,20,21,26,27,29,31,34,35,37,38,40,42</sup> 10 (33.3%) were bacterial,<sup>16,18,19,23,24,32,33,36,41,43</sup> (10, 33.3%), and 5 (16.6%) were other organism types<sup>22,25,28,30,39</sup> (Table 1 and Supplementary Table S3). In the clusters describing the composition of infected individuals, most were patients,<sup>14–21,23,24,26,27,29,31–38,40,41,43</sup> nurses,<sup>14,17,31,37</sup> advanced practice practitioners, and physicians,<sup>14,17,37</sup> and other healthcare workers.<sup>14,15,17,24,26,31,37,42,44</sup> Of 3,359 persons described in the clusters, 93.1% were patients and the remainder were healthcare workers (Supplementary Table S4).

### Data visualization types and elements

In total, 37 data visualizations were reported in the 30 eligible articles (Supplementary Table S3), with up to 2 visualizations per study. More than half of the data visualizations contained spatial elements,<sup>14–16,19,22–27,31,33–37,39–42,45</sup> and/or person type,<sup>14,15,17–22,24–32,34–36,39,40</sup> and half contained time elements.<sup>15–17,21,22,24–28,30,31,34,36,39,43</sup> Case type (confirmed, probable, or suspected)<sup>14,17,19,23,24,27,28,33,35,36,39,43</sup> and pathway type were less frequently incorporated.<sup>17,22,28–31,34,36</sup> None of the data visualizations contained contagious periods, and very few contained symptom-onset date (Table 2).<sup>14,36</sup>



**Fig. 2.** PRISMA-P flow diagram depicting articles identified through literature search, including those included in the analysis of data visualizations of transmission pathways in healthcare-associated infectious diseases outbreaks. Note. NHSN, National Healthcare Safety Network.

Timeline-type data visualizations were more likely to include person type<sup>15,24,26,28,36</sup> but more than half of timelines included spatial elements<sup>15,24,26</sup> and test positivity date.<sup>15,24,36</sup> Among floor-plan transmission maps,<sup>19,23,27,33,35,39–42</sup> time elements were only included in 20% of the studies.<sup>27,39</sup> All 3 cluster maps<sup>14,15,37</sup> prioritized spatial elements; however, 1 cluster map<sup>14</sup> successfully incorporated person type, case type, transmission pathways, symptom-onset date, and positive test date. Transmission trees<sup>20</sup> and social networks<sup>18</sup> successfully incorporated person types. For the heat map<sup>19</sup> and case proximity map<sup>16</sup> analyses, spatial elements were included. However, transmission networks<sup>17,22,25,27,29,30,32,34,36,38,43</sup> were more variable in the type of elements included to display transmission events; more than half contained time elements<sup>17,22,25,27,30,34,36,43</sup> and/or person types.<sup>17,22,25,27,29,30,32,34,36</sup>

Only 2 studies reported the software used to create data visualizations (2 [5.4%] of 37 data visualizations).<sup>18,19</sup> The following software programs were used to create data visualizations: AutoCAD,<sup>19</sup> QGIS,<sup>19</sup> and Pajek 4.90<sup>18</sup> (Supplementary Table S5).

## Discussion

In this scoping review of 1,957 articles, 30 studies with 37 data visualizations were identified that described clusters of infections in healthcare settings. Across the 37 data visualizations, more than half included spatial elements and patients, and half included time. Almost half contained case type and none included contagious periods. Overall, this scoping review has demonstrated the wide range of data visualization practices used in characterizing outbreaks. Although none of the data visualizations analyzed in this review incorporated all data elements (ie, contagious period, healthcare workers, pathways, symptom onset date, and test positive date), several types of visualizations incorporated multiple elements.

We observed significant variability the data elements included across visualization types. With the wide range of data elements included in each of the visualizations analyzed for this review, an infection preventionist should explore different data visualization types to analyze the specific circumstance of the cluster most effectively. This variability also suggests an opportunity to develop a

**Table 1.** Characteristics of 31 Studies Examining Clusters of Infectious Pathogens With Data Visualizations in Healthcare Settings

Study	Year	Country	Study Type	Setting	Scope of Outbreak	Infectious Pathogen	Cluster Duration	Cluster Size
Abbas et al	2021	Switzerland	Case-Control	Long-term care facility	Palliative care ward, rehabilitation care ward	SARS-CoV-2	35 d	37
Burns et al	2011	Ireland	Prospective	Acute care	Entire healthcare facility	Hepatitis B virus	425 d	7
Borges et al	2021	Portugal	Prospective	Acute care	Internal medicine, hemodialysis units	SARS-CoV-2	25 d	48
Charpentier et al	2017	France	Case-control	Tertiary care	Transplant unit, surgical cardiology outpatient clinic, echocardiography, transplant outpatient clinic, nephrology unit	<i>Pneumocystis jirovecii</i>	61 d	12
Cherifi et al	2006	Belgium	Prospective	Tertiary care	Geriatric ward	<i>Clostridioides difficile</i>	731 d	21
Da Silva et al	2021	Brazil	Retrospective	Acute care	Intensive and non-intensive care units	<i>Klebsiella pneumoniae</i>	365 d	24
Duong et al <sup>a</sup>	2016	Vietnam	Prospective	Acute care	Hemodialysis unit	Hepatitis C virus	720 d	11
Furusyo et al <sup>a</sup>	2004	Japan	Prospective	Acute care	Hemodialysis unit	Hepatitis C virus	5,204 d	12
Gandhi et al <sup>a</sup>	2013	South Africa	Retrospective	Acute care	General medicine ward	<i>Mycobacterium tuberculosis</i>	700 d	148
Jacobson et al	2015	South Africa	Prospective	Tertiary care	Hematology ICU	<i>Klebsiella pneumoniae</i>	30 d	8
Javaid et al	2021	USA	Prospective	Acute care	Psychiatric unit, oncology unit, emergency department, medical/surgical inpatient unit, cardiac interventional observation, inpatient telemetry unit, cardiac care unit, ICU, medical unit, cardiac catheter lab	Influenza A virus	12 d	107
Jia et al <sup>a</sup>	2016	China	Retrospective	Acute care	Cardiac surgery unit	<i>Clostridioides difficile</i>	792 d	20
Klompas et al <sup>a</sup>	2021	USA	Retrospective	Tertiary care	4 inpatient units	SARS-CoV-2	90 d	52
Kossow et al	2019	Germany	Prospective	Tertiary care	Obstetrics department, neonatal department	<i>Staphylococcus aureus</i>	401 d	8
Lee et al	2021	Korea	Retrospective	Acute care	Neurosurgery ICU, neurosurgery ward, radiology, bank, intervention room, echo lab, general internal medicine wards,	SARS-CoV-2	36 d	36
Marmor et al	2020	Australia	Retrospective	Tertiary care	Renal unit, surgical unit, ICU, hematology unit	Enterobacteriaceae	1,827 d	17
Moldovan et al	2019	Canada	Case-control	Tertiary care	3 campuses	<i>Staphylococcus aureus</i>	426 d	547
Nevez et al <sup>a</sup>	2018	France	Case-control	Acute care	Nephrology department	<i>Pneumocystis jirovecii</i>	402 d	25
Pagani et al	2015	Switzerland	Prospective	Geriatric hospital	Hematology ICU	Influenza virus	90 d	69
Pai et al	2020	USA	Retrospective	Acute care	Across inpatient units	<i>Clostridioides difficile</i>	2,526 d	1,963
Pérez-Lago et al	2021	Spain	Retrospective	Tertiary care	Gastroenterology ward	SARS-CoV-2	27 d	18
Rabodonirina et al	2004	France	Prospective	Tertiary care	Entire healthcare facility	<i>Pneumocystis jirovecii</i>	1,080 d	45
Shen et al	2003	China	Prospective	Acute care	Entire healthcare facility	SARS-CoV-1	4,621 d	77
Spada et al <sup>a</sup>	2008	Italy	Prospective	Acute care	Hemodialysis unit	Hepatitis C virus	184 d	14
Tsutsumi et al	2005	Japan	Retrospective	Geriatric hospital	Dementia ward	<i>Sarcoptes scabiei</i>	216 d	20
Varia et al	2003	Canada	Prospective	Community hospital	Entire healthcare facility	SARS-CoV-1	52 d	128
Vindrios et al	2017	France	Prospective	Acute care	Cardiac surgery unit	<i>Pneumocystis jirovecii</i>	185 d	7

(Continued)

**Table 1.** (Continued)

Study	Year	Country	Study Type	Setting	Scope of Outbreak	Infectious Pathogen	Cluster Duration	Cluster Size
Wee <i>et al</i>	2020	Singapore	Prospective	Tertiary care	Campus-wide	SARS-CoV-2	212 d	14
Wong <i>et al</i>	2010	China	Prospective	Tertiary care	General medicine ward	Influenza virus	16 d	9
Yang <i>et al</i> *	2010	Taiwan	Prospective	Community hospital	Respiratory care ward	<i>Haemophilus influenzae</i>	25 d	12

Note: End date of cluster in calculation (1 day was added).

\*Article did not describe the type of care provided at the healthcare facility so acute care was marked.

data visualization software tool that incorporates all the elements or has the capability to create different visualizations from the same cluster data set. The very infrequent reporting of software programs in these clusters did not allow us to evaluate whether a specific program could create hypothesized transmission pathways using different visualization types.

Among all data visualizations, 26 (70%) used spatial elements<sup>14–16,19,22–27,31,33–37,39–43</sup> and 19 (51%) included time elements.<sup>15–17,21,22,24–28,30,31,34,36,39,43</sup> The high frequency of these data element types may be related to accessibility and availability of information at a given healthcare facility. Less frequently utilized data visualization elements, such as contagious period and symptom-onset date, may relate to data accessibility. Despite the additional burden that collection of these elements place on healthcare facilities, symptom-onset dates and contagious periods should be collected and considered important features in cluster-transmission visualizations. For contact tracing and the development of hypotheses of transmission, these types of elements are essential to fully understanding the extent of the transmission event.

Specific data elements including patient information, microbiology cultures, as well as locations of patients and/or healthcare workers in space fit into the way in which standard outbreak investigations are conducted by infection preventionists. However, these data may not necessarily be in a form extractable directly from electronic health records (EHR). Thus, a standardized data visualization tool would need to be aware of the data infection preventionists have readily available. Additionally, the adoption of data visualization tools must include the consideration of data integration because EHR data may be difficult to extract and the software must work in tandem. EHR data mining has been effective at identifying transmission routes in healthcare outbreaks.<sup>46</sup> A data visualization tool that can be used to visualize and hypothesize transmission routes and that seamlessly integrates with the existing EHR could aid in informing interventions to stop transmission. However, data visualization tools need operational data integration to be effective.

Within the 37 data visualizations we analyzed, the frequency of spatial elements and time elements was higher than the 15% frequency of spatiotemporal elements reported among data visualizations in 2014.<sup>8</sup> Of the eligible articles, 19 of the 30 were published after 2014, which may account for some of the change. Of these 19 articles, 11 included spatial and time elements in data visualizations. The increased inclusion of spatiotemporal

elements may be an indicator of improving data visualization utility over time.

We observed a slight pattern in the types of visualizations used to analyze droplet versus contact transmission. Infection clusters due to bacteria, most commonly requiring contact precautions, used case-proximity maps, floor-plan transmission maps, and transmission networks more frequently. In contrast, clusters due to viruses, more likely requiring droplet precaution isolation (excepting hepatitis B and C viruses), used timelines, cluster maps, Gantt charts, and transmission networks most often. In addition, many of the viral cluster visualizations included patients and healthcare workers, whereas bacterial cluster visualizations mostly included patients. Inclusion of both healthcare workers and patients in viral cluster visualizations may be due to transmission patterns of respiratory pathogens such as SARS-CoV-2 and influenza A.<sup>47,48</sup> Data visualizations used for different infectious pathogens and the data elements used to describe the clusters could be tailored to the mechanisms of transmission of the pathogen under investigation.

Only 2 of the evaluated visualizations described the type of software used. Software such as R and GIS require training to use; other software created to aid in cluster visualization, such as DotMapper and HospMapper, require specific software knowledge and have technological limitations including data formatting (Supplementary Table S5). An ideal data visualization software program would not only be easily integrated into the current infection prevention workflow but would also be accessible for those with minimal programming expertise.

All the outbreaks in the articles were described after the outbreak occurred within the healthcare facility and after infection prevention measures were initiated. The data visualizations contained within these articles illustrated the transmission events and pathways that were already known. Therefore, the ability of these software tools to derive hypotheses about transmission pathways is limited, and this feature is essential when an outbreak is first identified to advance the investigation. Other limitations of this analysis include a small sample size of only 30 publications and 37 data visualizations, with most studies describing outbreaks in acute-care facilities. The common elements found in these articles may not be generalizable to other healthcare facilities seeking to visualize transmission pathways, such as long-term care facilities, behavioral health facilities, or psychiatric facilities. Our review may have been affected by publication bias toward acute-care facilities, which have better data accessibility, and by bias toward reporting larger outbreaks. Our findings may

**Table 2.** Types of Data Elements Included Across 37 Transmission Visualizations Grouped by Data Visualization Type in the 30 Eligible Articles Reviewed

Study	Infectious Pathogen	Data Visualization Type	Elements Included in the Data Visualization										
			Person Type	Patients	Healthcare workers	Other	Case Type	Pathway	Spatial Element	Time Element	Symptom Onset Date	Test Positive Date	Contagious Period
Javid et al	Influenza A virus	Timeline	✓	✓	✓	...	...	...	✓	✓	...	✓	...
Javid et al	Influenza A virus	Timeline	...	...	...	...	...	...	✓	✓	...	...	...
Jia et al	<i>Clostridioides difficile</i>	Timeline	✓	✓	...	...	...	...	...	✓	...	✓	...
Kossow et al	<i>Staphylococcus aureus</i>	Timeline	✓	✓	✓	...	✓	...	✓	✓	...	✓	...
Pagani et al	Influenza virus	Timeline	✓	✓	✓	...	✓	...	✓	✓	...	...	...
Rabodonirina et al	<i>Pneumocystis jirovecii</i>	Timeline	✓	✓	...	...	...	✓	...	✓	...	...	...
<b>Timeline Subtotal no. (%)</b>			5 (83)	5 (83)	3 (50)	...	2 (33)	1 (17)	4 (67)	6 (100)	...	3 (50)	...
Borges et al	SARS-CoV-2	Gantt chart	✓	...	✓	...	...	...	✓	✓	...	✓	...
Borges et al	SARS-CoV-2	gantt chart	✓	✓	✓	✓	...	...	✓	...	...	...	...
Burns et al	Hepatitis B virus	Gantt chart	✓	✓	...	...	...	...	...	✓	...	...	...
<b>Gantt Chart Subtotal, no. (%)</b>			3 (100)	3 (100)	2 (67)	1 (33)	...	...	...	...	...	1 (33)	...
Cherifi et al	<i>Clostridioides difficile</i>	Floor plan transmission map	...	...	...	...	✓	...	✓	...	...	✓	...
da Silva et al	<i>Klebsiella pneumoniae</i>	Floor plan transmission map	✓	✓	...	...	✓	...	✓	—	—	—	—
Furusyo et al	Hepatitis C Virus	Floor plan transmission map	✓	✓	...	...	✓	...	✓	...	...	...	...
Jacobson et al	<i>Klebsiella pneumoniae</i>	Floor plan transmission map	...	...	...	...	✓	...	✓	...	...	...	...
Pérez-Lago et al	SARS-CoV-2	Floor plan transmission map	✓	✓	...	...	✓	✓	✓	✓	...	✓	...
Tsutsumi et al	<i>Sarcoptes scabiei</i>	Floor plan transmission map	✓	✓	...	...	✓	...	✓	✓	...	...	...
Wong et al	Influenza A virus	Floor plan transmission map	✓	✓	...	...	...	...	✓	...	✓	...	...
Yang et al	<i>Haemophilus influenzae</i>	Floor plan transmission map	...	...	...	...	...	...	✓	...	...	...	...
Wee et al	SARS-CoV-2	Floor plan transmission map	...	...	...	...	...	...	✓	...	...	...	...
Lee	SARS-CoV-2	Floor plan transmission map	...	...	...	...	...	...	✓	...	...	...	...
<b>Floor plan transmission map subtotal, no. (%)</b>			5 (50)	5 (50)	...	...	6 (60)	1 (10)	10 (100)	2 (20)	1 (10)	2 (20)	...
Klompas et al	SARS-CoV-2	Cluster map	✓	...	✓	...	✓	✓	✓	...	...	✓	—
Lee et al	SARS-CoV-2	Cluster map	...	...	...	...	...	...	✓	...	...	...	...
Javid et al	Influenza A virus	Cluster map	✓	✓	✓	...	...	...	✓	...	...	...	...

(Continued)

**Table 2.** (Continued)

Study	Infectious Pathogen	Data Visualization Type	Elements Included in the Data Visualization										
			Person Type	Patients	Healthcare workers	Other	Case Type	Pathway	Spatial Element	Time Element	Symptom Onset Date	Test Positive Date	Contagious Period
<b>Cluster map subtotal, no. (%)</b>			2 (67)	2 (67)	2 (67)	...	1 (33)	1 (33)	3 (100)	—	1 (33)	1 (33)	—
Gandhi et al	<i>Mycobacterium tuberculosis</i>	Transmission network	✓	✓	...	...	...	✓	...	—	...	...	...
Marmor et al	Enterobacteriaceae	Transmission network	...	...	...	...	✓	...	✓	✓	...	...	...
Pérez-Lago et al	SARS-CoV-2	Transmission network	✓	✓	...	...	✓	✓	...	✓	...	...	...
Shen et al	SARS-CoV-1	Transmission network	...	...	...	...	...	...	...	—	...	...	...
Spada et al	Hepatitis C virus	Transmission network	✓	✓	...	...	...	...	...	—	...	...	...
Varia et al	SARS-CoV-1	Transmission network	✓	✓	✓	✓	✓	...	...	✓	...	...	...
Charpentier et al	<i>Pneumocystis jirovecii</i>	Transmission network	✓	✓	...	...	...	✓	✓	✓	...	...	...
Duong et al	Hepatitis C virus	Transmission network	✓	✓	...	...	...	...	✓	✓	—	—	—
Jia et al	<i>Clostridioides difficile</i>	Transmission network	✓	✓	...	...	✓	✓	✓	✓	—	—	—
Nevez et al	<i>Pneumocystis jirovecii</i>	Transmission network	✓	✓	...	...	✓	—	✓	✓	—	✓	—
Vindrios et al	<i>Pneumocystis jirovecii</i>	Transmission network	✓	✓	...	...	...	...	...	✓	—	✓	—
<b>Transmission network subtotal, no. (%)</b>			9 (82)	9 (82)	1 (9)	1 (9)	5 (45)	4 (36)	5 (45)	8(73)	—	2(18)	—
Abbas et al	SARS-CoV-2	Transmission tree	✓	✓	✓	✓	...	...	...	—	—	—	—
<b>Transmission tree subtotal, no. (%)</b>			1 (100)	1 (100)	1 (100)	1 (100)	...	...	...	...	...	—	—
Moldovan et al	<i>Staphylococcus aureus</i>	Social network <sup>a</sup>	✓	✓	✓	...	...	...	...	—	—	—	—
<b>Social network analysis subtotal, no.(%)</b>			1 (100)	1 (100)	1 (100)	...	...	...	...	...	...	—	—
Pai et al	<i>Clostridioides difficile</i>	Case proximity graph	...	...	...	...	...	...	✓	✓	—	—	—
<b>Spatial proximity map, no. (%)</b>			...	...	...	...	...	...	...	...	—	—	—
da Silva et al	...	...	...	...	...	...	...	...	✓	—	—	—	—
<b>Heat map subtotal, no. (%)</b>			...	...	...	...	...	...	...	...	—	—	—
<b>Totals N=37, no. (%)</b>			26 (70)	26 (70)	11 (30)	3 (8)	14 (39)	7 (19)	26 (70)	19(51)	2(5)	9(24)	—

<sup>a</sup>Data visualizations with specific software stated, see Supplementary Table 5.

underrepresent smaller outbreaks and investigations in lower-resourced facilities. Although the findings of this study included outbreaks reported internationally, limiting the search to English language and 1 database may have excluded less readily accessible and non-peer-reviewed data.

The data gathered from these articles will be used to inform the creation of a standardized visualization tool that can aid in current SARS-CoV-2 cluster investigations, with the overall goal of allowing for visualization of other infectious pathogen clusters in healthcare settings to reduce nosocomial transmission. Data visualization should include spatial elements, time elements, and elements that inform the exposure window and incubation period, as well as epidemiologic characteristics potentially associated with transmission such as healthcare worker type. No single data visualization may capture all informative elements, however. By understanding the types of common data visualization elements utilized in transmission visualizations, IPs can develop a generalized understanding of the types of possible data visualizations and data elements that can be included for different infectious pathogens. This study also highlights the need for a standardized data visualization that can utilize electronic medical records to allow for the generation of hypotheses regarding transmission and, therefore, improve public health measures and patient safety through faster infection prevention and control interventions to interrupt transmission.

**Acknowledgments.** We thank the UPMC Presbyterian-Shadyside Infection Prevention & Control team, as well as instructors and advisors at the University of Pittsburgh School of Public Health for their support on this manuscript.

**Financial support.** No financial support was provided relevant to this article.

**Conflicts of interest.** All authors report no potential conflicts of interest relevant to this article.

**Supplementary material.** To view supplementary material for this article, please visit <https://doi.org/10.1017/ash.2022.237>

## References

- Principles of epidemiology: lesson 6, section 2, self-study course SS1978. Centers for Disease Control and Prevention website. <https://www.cdc.gov/csels/dsepd/ss1978/lesson6/section2.html>. Accessed February 7, 2022.
- Rutter H, Parker S, Stahl-Timmins W, et al. Visualising SARS-CoV-2 transmission routes and mitigations. *BMJ* 2021;375:e065312.
- Sundermann AJ, Chen J, Kumar P, et al. Whole-genome sequencing surveillance and machine learning of the electronic health record for enhanced healthcare outbreak detection. *Clin Infect Dis* 2021. doi: [10.1093/cid/ciab946](https://doi.org/10.1093/cid/ciab946).
- Smith CM, Hayward AC. DotMapper: an Open Source tool for creating interactive disease point maps. *BMC Infect Dis* 2016;16:145.
- Smith CM, Allen DJ, Nawaz S, et al. An interactive data visualisation application to investigate nosocomial transmission of infections. *Wellcome Open Res* 2019;4:100.
- Salinas JL, Kritzman J, Kobayashi T, Edmond MB, Ince D, Diekema DJ. A primer on data visualization in infection prevention and antimicrobial stewardship. *Infect Control Hosp Epidemiol* 2020;41:948–957.
- Carroll LN, Au AP, Detwiler LT, Fu T-C, Painter IS, Abernethy NF. Visualization and analytics tools for infectious disease epidemiology: a systematic review. *J Biomed Inform* 2014;51:287–298.
- Davis GS, Sevdalis N, Drumright LN. Spatial and temporal analyses to investigate infectious disease transmission within healthcare settings. *J Hosp Infect* 2014;86:227–243.
- Schifman RB, Palmer RA. Surveillance of nosocomial infections by computer analysis of positive culture rates. *J Clin Microbiol* 1985;21:493–495.
- Grota PG, Stone PW, Jordan S, Pogorzelska M, Larson E. Electronic surveillance systems in infection prevention: organizational support, program characteristics, and user satisfaction. *Am J Infect Control* 2010;38:509–514.
- Dealing with outbreaks. World Health organization website. <https://www.who.int/csr/resources/publications/whodcscsreph200212.pdf>. Published 2002. Accessed November 24, 2021.
- NHSN organism list. Centers for Disease Control website. <https://www.cdc.gov/nhsn/pdfs/validation/2019/2019-NHSN-Organisms-List-Validation.xlsx>. Published 2019. Accessed October 13, 2021.
- Brady M, VonVille H, White J, et al. Transmission mapping of healthcare infection clusters: a scoping review protocol. Open Science Framework website. <https://osf.io/azuyr/>. Published August 11, 2021. Accessed March 13, 2022.
- Klompas M, Baker MA, Rhee C, et al. A SARS-CoV-2 cluster in an acute-care hospital. *Ann Intern Med* 2021;174:794–802.
- Javadi W, Ehni J, Gonzalez-Reiche AS, et al. Real-time investigation of a large nosocomial influenza: an outbreak informed by genomic epidemiology. *Clin Infect Dis* 2021;73:e4375–e4383.
- Pai S, Polgreen PM, Segre AM, Sewell DK, Pemmaraju SV, CDC MIND-Healthcare group. Spatiotemporal clustering of in-hospital *Clostridioides difficile* infection. *Infect Control Hosp Epidemiol* 2020;41:418–424.
- Varia M, Wilson S, Sarwal S, et al. Investigation of a nosocomial outbreak of severe acute respiratory syndrome (SARS) in Toronto, Canada. *CMAJ* 2003;169:285–292.
- Moldovan ID, Suh K, Liu EY, Jolly A. Network analysis of cases with methicillin-resistant *Staphylococcus aureus* and controls in a large tertiary-care facility. *Am J Infect Control* 2019;47:1420–1425.
- da Silva PP, da Silva FA, Rodrigues CAS, et al. Geographical information system and spatial-temporal statistics for monitoring infectious agents in hospital: a model using *Klebsiella pneumoniae* complex. *Antimicrob Resist Infect Control* 2021;10:92.
- Abbas M, Robalo Nunes T, Cori A, et al. Explosive nosocomial outbreak of SARS-CoV-2 in a rehabilitation clinic: the limits of genomics for outbreak reconstruction. *J Hosp Infect* 2021;117:124–134.
- Burns K, Heslin J, Crowley B, et al. Nosocomial outbreak of hepatitis B virus infection involving two hospitals in the Republic of Ireland. *J Hosp Infect* 2011;78:279–283.
- Charpentier E, Garnaud C, Wintenberger C, et al. Added value of next-generation sequencing for multilocus sequence typing analysis of a *Pneumocystis jirovecii* pneumonia outbreak. *Emerg Infect Dis* 2017;23:1237–1245.
- Cherifi S, Delmee M, Van Broeck J, Beyer I, Byl B, Mascart G. Management of an outbreak of *Clostridium difficile*-associated disease among geriatric patients. *Infect Control Hosp Epidemiol* 2006;27:1200–1205.
- Kossow A, Kampmeier S, Schaumburg F, Knaack D, Moellers M, Mellmann A. Whole-genome sequencing reveals a prolonged and spatially spread nosocomial outbreak of Pantone-Valentine leucocidin-positive methicillin-resistant *Staphylococcus aureus* (USA300). *J Hosp Infect* 2019;101:327–332.
- Nevez G, Le Gal S, Noel N, et al. Investigation of nosocomial pneumocystis infections: usefulness of longitudinal screening of epidemic and post-epidemic pneumocystis genotypes. *J Hosp Infect* 2018;99:332–345.
- Pagani L, Thomas Y, Huttner B, et al. Transmission and effect of multiple clusters of seasonal influenza in a Swiss geriatric hospital. *J Am Geriatr Soc* 2015;63:739–744.
- Pérez-Lago L, Martínez-Lozano H, Pajares-Díaz JA, et al. Overlapping of independent SARS-CoV-2 nosocomial transmissions in a complex outbreak. *mSphere* 2021;6(4):e0038921.
- Rabodonirina M, Vanhems P, Couray-Targe S, et al. Molecular evidence of interhuman transmission of *Pneumocystis pneumonia* among renal transplant recipients hospitalized with HIV-infected patients. *Emerg Infect Dis* 2004;10:1766–1773.
- Spada E, Abbate I, Sicurezza E, et al. Molecular epidemiology of a hepatitis C virus outbreak in a hemodialysis unit in Italy. *J Med Virol* 2008;80:261–267.
- Vindrios W, Argy N, Le Gal S, et al. Outbreak of *Pneumocystis jirovecii* Infection among heart transplant recipients: molecular investigation and management of an interhuman transmission. *Clin Infect Dis* 2017;65:1120–1126.

31. Borges V, Isidro J, Macedo F, *et al*. Nosocomial outbreak of SARS-CoV-2 in a “Non-COVID-19” hospital ward: virus genome sequencing as a key tool to understand cryptic transmission. *Viruses* 2021;13(4):604. doi: [10.3390/v13040604](https://doi.org/10.3390/v13040604).
32. Gandhi NR, Weissman D, Moodley P, *et al*. Nosocomial transmission of extensively drug-resistant tuberculosis in a rural hospital in South Africa. *J Infect Dis* 2013;207:9–17.
33. Jacobson RK, Manesen MR, Moodley C, *et al*. Molecular characterisation and epidemiological investigation of an outbreak of blaOXA-181 carbapenemase-producing isolates of *Klebsiella pneumoniae* in South Africa. *S Afr Med J* 2015;105:1030–1035.
34. Duong CM, McLaws M-L. An investigation of an outbreak of hepatitis C virus infections in a low-resourced hemodialysis unit in Vietnam. *Am J Infect Control* 2016;44:560–566.
35. Furusyo N, Kubo N, Nakashima H, Kashiwagi K, Etoh Y, Hayashi J. Confirmation of nosocomial hepatitis C virus infection in a hemodialysis unit. *Infect Control Hosp Epidemiol* 2004;25:584–590.
36. Jia H, Du P, Yang H, *et al*. Nosocomial transmission of *Clostridium difficile* ribotype 027 in a Chinese hospital, 2012–2014, traced by whole-genome sequencing. *BMC Genomics* 2016;17:405.
37. Lee U, Kim SE, Lee SY, *et al*. Source analysis and effective control of a COVID-19 outbreak in a university teaching hospital during a period of increasing community prevalence of COVID-19. *J Korean Med Sci* 2021; 36(24):e179.
38. Shen Z, Ning F, Zhou W, *et al*. Superspreading SARS events, Beijing, 2003. *Emerging Infect Dis* 2004;10:256–260.
39. Tsutsumi M, Nishiura H, Kobayashi T. Dementia-specific risks of scabies: retrospective epidemiologic analysis of an unveiled nosocomial outbreak in Japan from 1989–90. *BMC Infect Dis* 2005;5:85.
40. Wong BCK, Lee N, Li Y, *et al*. Possible role of aerosol transmission in a hospital outbreak of influenza. *Clin Infect Dis* 2010;51:1176–1183.
41. Yang CJ, Chen TC, Wang CS, *et al*. Nosocomial outbreak of biotype I, multidrug-resistant, serologically non-typeable *Haemophilus influenzae* in a respiratory care ward in Taiwan. *J Hosp Infect* 2010;74:406–409.
42. Wee LE, Sim XYJ, Conceicao EP, *et al*. Containment of COVID-19 cases among healthcare workers: the role of surveillance, early detection, and outbreak management. *Infect Control Hosp Epidemiol* 2020;41: 765–771.
43. Marmor A, Daveson K, Harley D, Coatsworth N, Kennedy K. Two carbapenemase-producing Enterobacteriaceae outbreaks detected retrospectively by whole-genome sequencing at an Australian tertiary hospital. *Infect Dis Health* 2020;25:30–33.
44. Abbas M, Robalo Nunes T, Martischang R, *et al*. Nosocomial transmission and outbreaks of coronavirus disease 2019: the need to protect both patients and healthcare workers. *Antimicrob Resist Infect Control* 2021;10:7.
45. Memish ZA, Al-Tawfiq JA, Alhakeem RF, *et al*. Middle East respiratory syndrome coronavirus (MERS-CoV): a cluster analysis with implications for global management of suspected cases. *Travel Med Infect Dis* 2015;13: 311–314.
46. Sundermann AJ, Miller JK, Marsh JW, *et al*. Automated data mining of the electronic health record for investigation of healthcare-associated outbreaks. *Infect Control Hosp Epidemiol* 2019;40:314–319.
47. Sepkowitz KA. Occupationally acquired infections in healthcare workers. Part I. *Ann Intern Med* 1996;125:826–834.
48. Sepkowitz KA. Occupationally acquired infections in healthcare workers. Part II. *Ann Intern Med* 1996;125:917–928.
49. Ohannessian R, Bénét T, Argaud L, *et al*. Heat map for data visualization in infection control epidemiology: an application describing the relationship between hospital-acquired infections, Simplified Acute Physiological Score II, and length of stay in adult intensive care units. *Am J Infect Control* 2017;45:746–749.
50. Visualization Types–Data Visualization–LibGuides at Duke University. Duke University Library website. [https://guides.library.duke.edu/datavis/vis\\_types](https://guides.library.duke.edu/datavis/vis_types). Accessed November 8, 2021.