Journal of Refugee & Global Health

Original Research

Designing the Arriving Refugee Informatics Surveillance and Epidemiology (ARIVE) System: A Web-based Electronic Database for Epidemiological Surveillance

William A Mattingly[1], Ruth M Carrico[1], Timothy L Wiemken[1], Rahel S Bosson[1], Stephen P Furmanek, Robert R Kelley[1], Rebecca A Ford[2], Kimberley Buckner, Julio A Ramirez[1]*

Abstract

Objectives: This article discusses the design and implementation of the *Arriving Refugee Informatics surVeillance and Epidemiology (ARIVE)* system. The system seeks to improve the health of refugees undergoing resettlement and enhance existing health surveillance networks.

Methods: Using the REDCap Electronic Data Capture (EDC) software as a basis we create a refugee health database incorporating data from the Centers for Disease Control and Prevention's Electronic Disease Notification (EDN) system and domestic screening data from refugee health care providers.

Results: Domestic screening and EDN refugee health data have been integrated for 16,739 refugees resettled from 35 different countries into the state of Kentucky from the years 2013-2016.

Discussion: The ARIVE system is a flexible software system that implements the core of a health surveillance network in a way that is sustainable and cost-effective, and its data dictionary provides an easy way to share and improve the database structure of a health surveillance network.

DOI: 10.18297/rgh/vol2/iss2/6
Submitted Date: November 9, 2018
Accepted Date: July 11, 2019
Website: https://ir.library.louisville.edu/rgh
Affiliations:
[1]Department of Medicine, University of Louisville, Louisville, USA
[2]Kentucky Office of Refugees

Recommended Citation:
Mattingly, William A.; Carrico, Ruth M.; Wiemken, Timothy L.; Bosson, Rahel S.; Furmanek, Stephen P.; Kelley, Robert R.; Ford, Rebecca A.; Buckner, Kimberley; Ramirez, Julio A. (2019) "Designing the Arriving Refugee Informatics Surveillance and Epidemiology (ARIVE) System: A Web-based Electronic Database for Epidemiological Surveillance," *Journal of Refugee & Global Health*: Vol. 2 : Iss. 2, Article 6.

Introduction

Each year, the United States resettles thousands of refugees through the U.S. Refugee Program, with numbers of refugees cared for in US communities surging well past 140,000 during fiscal year 2015[1, 2]. As these refugees enter the United States from a variety of countries, they bring with them communicable and non-communicable health conditions. These conditions represent both short and long-term challenges to public and personal health. Surveillance of health conditions that are impactful to public health as well as those that may require high resource utilization, like chronic kidney disease, is critical to the resettlement program as well as to the communities that welcome the refugees.

Prior to 2008, health information related to immigrants and refugees entering the United States was managed using a paper process. In October 2008, the Electronic Disease Notification (EDN) System developed by the Centers for Disease Control became the primary system for tracking immigrants and refugees with Class A and B medical conditions or those who possessed HIV. The EDN has been effective in providing disease information access to domestic health agencies in a timely manner, but the health surveillance of refugees and immigrants after resettlement is impeded by the absence of a consistent refugee or immigrant status field in most electronic health records. This means there are little data available to evaluate the effectiveness of overseas medical examinations prior to resettlement or to establish evidence-based guidelines for post-arrival medical examination.

In 2012, the Kentucky Office for Refugees in partnership with the University of Louisville School of Medicine, began development of a process to integrate and store electronic data gathered from overseas medical examinations with data gathered during domestic health screenings of refugee resettlement in Kentucky. Reports to the CDC describing the state of refugee health in Kentucky can then be generated electronically, addressing the need for timely reporting and notification of infectious diseases as well as chronic health conditions. This electronic process also provides a more complete picture of the health status of individual refugees in Kentucky, potentially improving refugee quality of care and long-term health outcomes. In this article, we describe the development and impact of the *Arriving*

***Correspondence To:**
William A. Mattingly, PhD
501 E Broadway, Louisville, KY 40202
(502) 852-1555

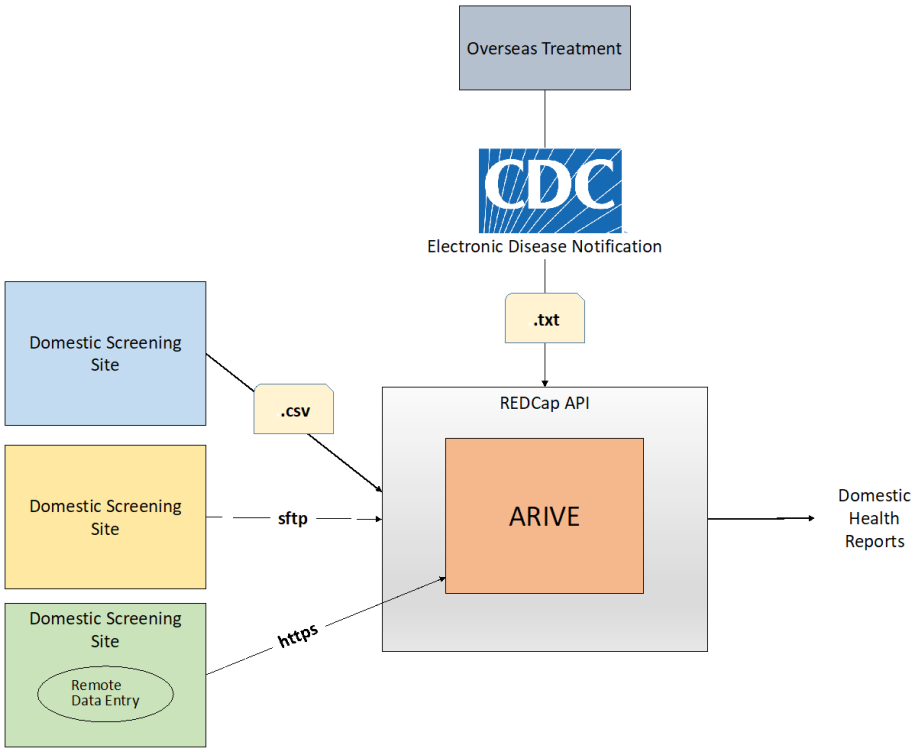

Figure 1 The ARIVE system incorporates refugee health data from multiple sources including overseas health assessment and domestic health screening. Overseas assessments are stored and available through the Center for Disease Control and Prevention's Electronic Disease Notification (EDN) system. REDCap's Application Programmer Interface serves as the bridge to import this data into ARIVE. Domestic clinics performing refugee health screenings share their screening data with the ARIVE team. Depending on the data capabilities of the clinic, file upload, data entry or automated connection are used to transfer and incorporate data. ARIVE then generates aggregate reports for the CDC and the state office for refugees and other stakeholders.

Refugee Informatics surVeillance and Epidemiology (ARIVE) system, and discuss its application to provide public health surveillance for refugee resettlement in Kentucky.

Methods

System Design

Combining EDN data with domestic screening data required collaboration between the ARIVE team at the University of Louisville, clinic sites providing care for refugees, and the Kentucky office for refugees. To facilitate combining this data for surveillance, an observational study was approved by the university institutional review board with IRB number 16.0350 and business associates agreements were formed between the Kentucky Office for Refugees, the university and each of the eight collaborating screening sites. Multiple methods of data transfer are necessary because of the differences in data management at each site. Some sites had paper forms to share while for others a connection to an sFTP repository was possible from the screening Electronic Health Record (EHR) system.

Software

We used previously designed health surveillance systems as models for the development of ARIVE.[3-7] These included systems optimized for use in limited resource environments[8-11], developing countries[12-18], and rural settings of developed countries[19]. In terms of licensing costs, free and open-source software provides the most affordable solution. Notable freely available medical database software includes electronic medical records like OpenMRS[20] and electronic data capture solutions like OpenClinica and REDCap[21]. REDCap's strong community of users and previous successes at being adapted and used for health-related projects formed the basis for using it as the core software component of the ARIVE system. REDCap provides the many security features needed for HIPAA compliance such as secure password access and automatic logout of screens with sensitive data. The REDCap software is free to register and download for use on work that is not for profit.[22]

Figure 1 shows the layout for the ARIVE system. Overseas assessments are stored and available through the Centers for Disease Control and Prevention's Electronic Disease Notification (EDN) system. REDCap's Application Programmer Interface serves as the bridge to import these data into ARIVE. Domestic clinics performing refugee health screenings share their screening data with the ARIVE team and depending on the data capabilities of the clinic, file upload, data entry or automated connection are used to transfer and incorporate data. ARIVE then generates aggregate reports for the CDC, the state office for refugees, and other stakeholders.

Environment

The software environment for the ARIVE system is a virtual Linux server hosted in the university's enterprise datacenter. This server environment affords the necessary physical requirements for storing data covered by HIPAA, namely, secured physical access only available to authorized datacenter personnel. Requirements for installation included a Linux based operating system, a web server supporting the PHP server-side scripting language, and a MySQL database store. This popular server environment is commonly referred to as LAMP (Linux-Apache-MySQL-PHP) and the ARIVE system uses version 6.8 of Red Hat Enterprise Linux.

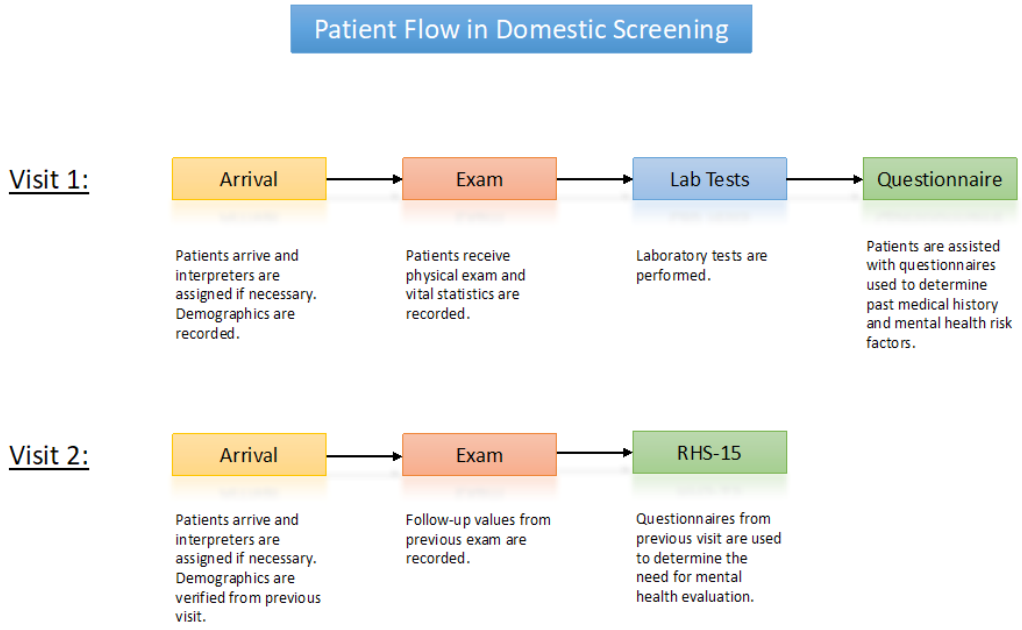

Figure 2 Screening process for arriving refugees. Domestic screening usually consists of two visits. The first visit includes patient exam and routine labs, followed by an in-depth questionnaire assisted by interpreters when necessary. The second visit includes a follow-up exam and additional questionnaires or lab tests depending on results from the first visit.

Database Design

The CDC guidelines for domestic refugee health screening were used to develop the data collection instruments and the corresponding data dictionary.[23, 24] An overview of domestic health screening is shown in **Figure 2**. Additionally, the screening instruments were reviewed and revised by users at the eight clinics performing domestic screening. The data collection instruments for ARIVE are shown in **Table 1**.

Results

Data Collection

Since 2012 information for 16,739 resettlements from 35 countries have been entered into ARIVE. After a 3 month period of design and development of the ARIVE system, we began the process of data collection. Firstly, providers at clinic sites providing care for refugees, or their designees, documented results of the initial domestic health screen for adults and children. An input screen for the demographics page of ARIVE is shown in Figure 3. When possible, connections were established between the EHR of the clinic and ARIVE. Otherwise, the site had the option of exporting reports and securely sending the files to the ARIVE team. Sites with low refugee volume were given a secure ARIVE account to remotely upload refugee data. Information available from prior health assessments including data from the EDN were added to domestic health screen data by the ARIVE team. Design of the database ensured that there was no overlap between the EDN set of variables and screening variables, preventing the possibility of accidental overwrite of information.

Routine data quality audits were performed and the most prevalent data quality issue was missing data. Most often this was attributed to laboratory results not completed at time of data entry. Legibility of data on the paper data collection form along with transposition of letters and numbers on data collection forms led to errors in data entry. Periodic follow-up with screening sites to provide corrected data resolved the majority of missing data problems. When this was not possible a flag would be set in a record indicating the missing data field and the types of correction and follow-up that were attempted.

Use

The ARIVE system has four main types of user roles: Data Entry, Investigator, Statistician and Data Administrator. Data entry users are responsible for entering data either remotely or on-site. This role is restricted to viewing, creating and updating records. The Investigator role is intended for collaborators to use ARIVE data for an approved study. This role allows creating comprehensive reports in addition to the creation and update of records. The role of statistician provides the additional capability of exporting the dataset for analysis and requesting a secure software key that can be used to access the data with a statistical programming language. Finally, data administrators have full access to all features of the system, including the assignment of user roles and modification of variable names and project structure. Two data entry users were from external sites outside the ARIVE team and their data access was customized to limit update access to data from their own institution. REDCap's user management system features account suspension and expiration features for data entry personnel or collaborators which leave the project. This feature allows long term logs of patient information access to be maintained in the database.

Reporting and Analysis

Report generation is a permission that is assigned to investigators, statisticians, and data administrators. Created reports can be restricted to only allow certain users to generate and run them. The ARIVE team have developed over 40 dynamic reports using REDCap's reporting syntax. A report displaying all records with a positive RHS-15 score since March 1, 2015 has the following reporting syntax: ([dov] >= "2015-03-01") AND ([rhs15_score] = "1")

Table 1 ARIVE data collection instruments

Instrument Name	Description	No. of Fields
Demographics	Patient's alien identification number, resettlement clinic, etc	26
Overseas Medical Documents	Class A status, Class B TB status, Class B Other status	18
Overseas Immunizations	Overseas Immunization information for MMR, TB, etc.	77
Immunizations	Domestic immunization information	122
Nursing Workup	Interpreter need and assignment, Allergy/Drug Use Questions	42
TB Questionnaire	TB symptom questionnaire	15
Vital Signs Day 1	Height, weight, BMI, temperature, etc.	23
Laboratory	Glucose, protein, blood, etc.	68
Intestinal Parasite Screening	Results of parasite screening	6
Additional Lab Results	If requested, results for chlamydia, gonorrhea, etc.	17
Tuberculosis Screening	X-ray and treatment history for TB	13
Visit 2	Follow-up labs and vitals for visit 2	27
Social Ethnographic History	Survey questions about family members/profession questions	48
Mental Health Assessment RHS 15	Standard mental health assessment	6
Past Medical History	Self-reported medical history	43
Review Of Systems	Symptoms grouped by body system	31
Physical Exam	Assessment of physical features and pain	59
Development Assessment	Development assessment for children under 18	25

The results of reports can be rendered as bar graphs, pie graphs, or scatter plots for suitable variables of interest. Routine reports to the Kentucky Office for Refugees (KOR) and CDC are created and exported using the reporting tool.

Collaborators may perform retrospective studies with ARIVE data upon approval by the Institutional Review Board. Several features of REDCap have been designed to make data export and analysis as convenient as possible. These features form a group of interfaces which is referred to as the REDCap application programmer interface (API). The REDCap API provides analysts with a common way to connect to its database despite using different programming languages. REDCap currently supports data export formats for SAS, SPSS, Stata, and R as well as real-time connections for the Python, Perl, R, and C programming languages.

Several published studies have focused on the data gathered in the ARIVE system[25, 26], including vaccine effectiveness in refugee populations[27], the role of refugees in nursing and physician resident curriculum[28], and refugee utilization of emergency room services[29]. Collaborators continue to make use of the ARIVE data to investigate research questions regarding refugee health in Kentucky.

Data Sharing

The design of the ARIVE database was intended to make sharing data with other organizations and institutions as straightforward as possible. Some of this can be done with the generation of reports and exports, but it is also possible to make direct connections from the ARIVE database to external systems, and the ARIVE system currently has two outgoing data connections facilitated by business associate and data use agreements.

The Kentucky Immunization Registry (KYIR) is known as an Immunization Information System (IIS). It provides deduplication and various data quality management tools to combine all the information for a particular individual into a single, accurate record. Once EDN immunization data is imported into ARIVE, it can be communicated to the state immunization registry as an HL7 file via a Minimal Lower Layer Protocol (MLLP) connection. Establishing this connection requires an exchange of secure socket layer (SSL) credentials between the state health information exchange and the ARIVE team. Using an HL7 immunization message template we create a REDCap API script to save ARIVE immunization information as HL7 files for periodic communication to the registry. Immunization records for 13,825 refugees are currently searchable in the state immunization registry.

ARIVE is also a member of the multi-state refugee surveillance network the Colorado Center for Excellence in Refugee Health. The Colorado Department of Public Health and Environment (CDPHE) provides a secure sFTP repository for each network member to upload de-identified records which are combined to show health trends of resettled refugees. The ARIVE team entered into a data use agreement with the CDPHE and a plugin with the provided sFTP credentials is used to periodically deposit de-identified records to the surveillance network. A total of 14,653 records have been shared to the network which includes all cases through the end of 2017.

Discussion

Differences in the screening process of sites initially caused gaps in the data points collected. These were resolved with in person and virtual meetings to standardize the screening process across each of the sites. Small changes continued to be made in the first couple of years of the project before the data collection became uniform.

The reporting function of the software made it much easier to manage the task of following up on missing data from sites. The majority of reports where created and run by the ARIVE team, but some were designed to be used by participating sites. This allowed sites using paper data collection instruments to see

printed reports of their data in aggregate.

The use of the system improved the general understanding of electronic data systems and disease surveillance among participating sites. Using uniform date formats and uniform naming conventions allowed the surveillance process to be standardized. This was new process for many collaborators involved in data collection.

Because of the expense of data entry personnel and limited funding available for new projects, automated systems still represent the most effective way to perform health surveillance. The primary expense for automated systems involve the high expertise in setting up, maintaining and migrating connections. Unfortunately, we found few sites that could overcome the resource barriers needed to commit to these requirements.

Conclusion

A flexible solution like the ARIVE system provided a stable and compliant way to achieve effective disease surveillance in current healthcare environments. It provides the capability to report disease and health trends to central and local health organizations, and the software used is available for free to non-profit institutions. As the results of the process are made available to other states and departments, they could become part of a uniform process of refugee health and disease surveillance in the United States.

Acknowledgements

Funding for this project was provided by CDC grants 1U50CK000288 and 3U50CK000288-04S1 and the Kentucky Office of Refugees. Partner sites for the project include Catholic Charities of Louisville, Louisville, KY; Family Health Centers – Americana, Louisville, KY; Bluegrass Community Health Center, Lexington, KY; Shawnee Community Health Center, Louisville, KY; Home of the Innocents, Louisville, KY; Fairview Community Health Center, Bowling Green, KY; Green River District Health Department, Owensboro, KY; ZipClinic Urgent Care, Bowling Green, KY. Thanks to the Kentucky Health Information Exchange (KYHIE) and the Kentucky Immunization Registry (KYIR) teams for their support. Funding for Colorado Center for Excellence in Refugee Health provided by CDC grant NU50CK000475.

References

1. Office of the Administration for Children & Families Office of Refugee Resettlement. Fiscal Year 2014 Refugee Arrivals. Office of Refugee Resettlement Resource Library. 2015 Feb 11. https://www.acf.hhs.gov/orr/resource/fiscal-year-2014-refugee-arrivals
2. Office of the Administration for Children & Families Office of Refugee Resettlement. FY 2015 Served Populations by State and Country of Origin. Office of Refugee Resettlement Resource Library. 2016 Apr 22. https://www.acf.hhs.gov/orr/resource/fy-2015-served-populations-by-state-and-country-of-origin-refugees-only
3. Reis BY, Kirby C, Hadden LE, Olson K, McMurry AJ, Daniel JB, et al. AEGIS: a robust and scalable real-time public health surveillance system. J Am Med Inform Assoc. 2007 Sep-Oct;14(5):581–8.
4. Cakici B, Hebing K, Grünewald M, Saretok P, Hulth A. CASE: a framework for computer supported outbreak detection. BMC Med Inform Decis Mak. 2010 Mar;10(1):14.
5. Krause G, Altmann D, Faensen D, et al. SurvNet electronic surveillance system for infectious disease outbreaks. Germany; 2007. https://doi.org/10.3201/eid1310.070253.
6. McMurry AJ, Gilbert CA, Reis BY, Chueh HC, Kohane IS, Mandl KD. A self-scaling, distributed information architecture for public health, research, and clinical care. J Am Med Inform Assoc. 2007 Jul-Aug;14(4):527–33.
7. Rolfhamre P, Jansson A, Arneborn M, Ekdahl K. SmiNet-2: description of an internet-based surveillance system for communicable diseases in Sweden. Euro Surveill. 2006;11(5):103–7. https://doi.org/10.2807/esm.11.05.00626-en PMID:16757847
8. Yi Q, Hoskins RE, Hillringhouse EA, Sorensen SS, Oberle MW, Fuller SS, et al. Integrating open-source technologies to build low-cost information systems for improved access to public health data. Int J Health Geogr. 2008 Jun;7(1):29.
9. Natter MD, Quan J, Ortiz DM, Bousvaros A, Ilowite NT, Inman CJ, et al. An i2b2-based, generalizable, open source, self-scaling chronic disease registry. J Am Med Inform Assoc. 2013 Jan;20(1):172–9.
10. Campbell TC, Hodanics CJ, Babin SM, Poku AM, Wojcik RA, Skora JF, et al. Developing open source, self-contained disease surveillance software applications for use in resource-limited settings. BMC Med Inform Decis Mak. 2012 Sep;12(1):99.
11. Pittman CA, Miranpuri AS. Neurosurgery clinical registry data collection utilizing Informatics for Integrating Biology and the Bedside and electronic health records at the University of Rochester. Neurosurg Focus. 2015 Dec;39(6):E16.
12. Yan WR, Nie SF, Xu B, Dong HJ, Palm L, Diwan VK. Establishing a web-based integrated surveillance system for early detection of infectious disease epidemic in rural China: a field experimental study. BMC Med Inform Decis Mak. 2012 Feb;12(1):4.
13. Rajput ZA, Mbugua S, Amadi D, Chepngeno V, Saleem JJ, Anokwa Y, et al. Evaluation of an Android-based mHealth system for population surveillance in developing countries. J Am Med Inform Assoc. 2012 Jul-Aug;19(4):655–9.
14. Steiner A, Hella J, Grüninger S, Mhalu G, Mhimbira F, Cercamondi CI, et al. Managing research and surveillance projects in real-time with a novel open-source eManagement tool designed for under-resourced countries. J Am Med Inform Assoc. 2016 Sep;23(5):916–23.
15. Tuti T, Bitok M, Malla L, Paton C, Muinga N, Gathara D, et al. Improving documentation of clinical care within a clinical information network: an essential initial step in efforts to understand and improve care in Kenyan hospitals. BMJ Glob Health. 2016 May;1(1):e000028.
16. Amboko BI, Ayieko P, Ogero M, Julius T, Irimu G, English M; Clinical Information Network authors. Malaria investigation and treatment of children admitted to county hospitals in western Kenya. Malar J. 2016 Oct;15(1):506.
17. Mbevi G, Ayieko P, Irimu G, Akech S, English M; Clinical Information Network authors. Prevalence, aetiology, treatment and outcomes of shock in children admitted to Kenyan hospitals. BMC Med. 2016 Nov;14(1):184.

18. Ayieko P, Ogero M, Makone B on behalf of the Clinical Information Network authors, et al. Characteristics of admissions and variations in the use of basic investigations, treatments and outcomes in Kenyan hospitals within a new Clinical Information Network. Archives of Disease in Childhood 2016;101:223-229.
19. Sheikhali SA, Abdallat M, Mabdalla S, Al Qaseer B, Khorma R, Malik M, et al. Design and implementation of a national public health surveillance system in Jordan. Int J Med Inform. 2016 Apr;88:58–61.
20. Seebregts CJ, Mamlin BW, Biondich PG, Fraser HS, Wolfe BA, Jazayeri D, et al.; OpenMRS Implementers Network. The OpenMRS implementers network. Int J Med Inform. 2009 Nov;78(11):711–20.
21. Tuti T, Bitok M, Paton C, Makone B, Malla L, Muinga N, et al. Innovating to enhance clinical data management using non-commercial and open source solutions across a multi-center network supporting inpatient pediatric care and research in Kenya. J Am Med Inform Assoc. 2016 Jan;23(1):184–92.
22. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009 Apr;42(2):377–81.
23. Division of Global Migration and Quarantine. General Refugee Health Guidelines. U.S. Department of Health and Human Services Centers for Disease Control and Prevention. 2012 Aug 6. https://www.cdc.gov/immigrantrefugeehealth/pdf/general.pdf
24. Division of Global Migration and Quarantine. Guidelines and Discussion of the History and Physical Examination. U.S. Department of Health and Human Services Centers for Disease Control and Prevention. 2012 Apr 16. https://www.cdc.gov/immigrantrefugeehealth/pdf/guidelines-history-physical.pdf
25. Bosson R, Carrico RM, Raghuram A, Peyrani P, Mattingly WA, Ford R, et al. Refugee-Centered Medical Home: A New Approach to Care at the University of Louisville Global Health Center. J Refug Glob Health. 2017;1(1):3.
26. Carrico RM, Wiemken TL, Bosson R, Mattingly WA, Ford R, Van Heiden S, et al. Health Profiles of Newly Arriving Refugees in Kentucky, 2016: Data from the University of Louisville Global Health Program. J Refug Glob Health. 2017;1(1):2.
27. Carrico RM, Goss L, Wiemken TL, Bosson RS, Peyrani P, Mattingly WA, et al. Infection prevention and control and the refugee population: Experiences from the University of Louisville Global Health Center. Am J Infect Control. 2017 Jun;45(6):673–6.
28. Carrico R, Bosson R, Koch J, Raghuram A, Peyrani P, Ford R, et al. Addressing the challenges of refugee health: experiences from the University of Louisville interprofessional refugee immunization clinic. J Nurs Educ Pract. 2015;5(12):88.
29. Carrico RM, Mitchell C, Balcom D, Goss L, Persaud A, Mutsch K, et al. Use of Emergency Department for Care Access by Refugees Resettling in Kentucky, 2015: Findings from the University of Louisville Global Health Center. J Refug Glob Health. 2017;1(1):5.