VIEWPOINT

The GapMap project: a mobile surveillance system to map diagnosed autism cases and gaps in autism services globally

Jena Daniels[1,2], Jessey Schwartz[1,2], Nikhila Albert[1,2,3], Michael Du[1,2] and Dennis P. Wall[1,2,4*]

Abstract

Although the number of autism diagnoses is on the rise, we have no evidence-based tracking of size and severity of gaps in access to autism-related resources, nor do we have methods to geographically triangulate the locations of the widest gaps in either the US or elsewhere across the globe. To combat these related issues of (1) mapping diagnosed cases of autism and (2) quantifying gaps in access to key intervention services, we have constructed a crowd-based mobile platform called "GapMap" (http://gapmap.stanford.edu) for real-time tracking of autism prevalence and autism-related resources that can be accessed from any mobile device with cellular or wireless connectivity. Now in beta, our aim is for this Android/iOS compatible mobile tool to simultaneously crowd-enroll the massive and growing community of families with autism to capture geographic, diagnostic, and resource usage information while automatically computing prevalence at granular geographical scales to yield a more complete and dynamic understanding of autism resource epidemiology.

Keywords: Autism, Autism spectrum disorder, Crowdsourcing, Prevalence, Resources, Epidemiology

Background

Across the globe, families affected by autism must navigate substantial gaps in availability of health care resources. The incidence of autism has been on the rise in the USA and across the globe, but the rate and exact prevalence of autism remain unclear. Estimates from the Centers for Disease Control (CDC) represent extrapolations from a small collection of states, and by the time of publication, they can be up to 2 years out of date. Today's alarming estimate of 1 in 68 children is based on data from 2014 [1, 2]. Actual rates are likely to be even higher, but until we have fully federated electronic medical records that are accessible on a national scale, we will not have a robust understanding of the true prevalence in the US let alone on a global scale for this lifelong condition.

As the number of autism diagnoses has increased, so too has the number of children with a high risk for autism that remain undiagnosed [3]. This translates into a waiting period estimated to average 13 months for a diagnostic assessment in the USA [4]. These wait times are even longer in rural environments and underserved areas with lower socioeconomic status [3–7]. Disparities in geographic coverage of resources often coincide with a lack of public awareness of autism spectrum disorders (ASD) and difficulty in finding diagnostic or post-diagnosis options with appropriate insurance coverage, exacerbating diagnostic delays and increasing the amount of hassle and stress for families navigating complexities associated with an ASD diagnosis throughout that individual's life. The resulting delays in diagnoses are also detrimental to children who are unable to receive behavioral and intervention therapies during the critical periods in which they are maximally impactful [8–12]. Frustratingly, however, we do not yet know the exact size and severity of gaps in access, where the widest divides exist globally, and the specific types of disparities across the life course as individuals with autism transition to adulthood.

Main text

Little attention has been paid to creating curated, easily Web-searchable, and comprehensive lists of autism

* Correspondence: dpwall@stanford.edu
[1]Department of Pediatrics, Division of Systems Medicine, Stanford University, Stanford, CA, USA
[2]Department of Biomedical Data Science, Stanford University, Stanford, CA, USA
Full list of author information is available at the end of the article

Daniels *et al. Molecular Autism* (2017) 8:55

Page 2 of 6

services. The few exceptions include Autism Speaks [13] and Autism Source [14], yet these account for only about 1500 unique resources, likely a fraction of the actual.

Resource gaps

Resource gaps (regions in which there exist limited diagnostic or treatment resources with respect to the demand) require comprehensive knowledge of both autism epidemiology and the geographic distribution of autism resources. Finding and understanding these resource gaps can drive novel innovation of products that can mobilize to the home and/or create shifts in resource usage to direct jobs and care towards particularly fillable gaps in care management for individuals with autism. This can be done through collecting robust hard data, allocating resources more efficiently, and providing information to emerging organizations and businesses to let them know where their services are needed most.

While autism epidemiology is a common research area [15] and Rzhetsky et al. found that the incidence of ASD is affected by the state-level regulatory and environmental factors [16], we are still far from understanding the true prevalence of diagnosed cases of autism. Many current epidemiological studies suffer from small sample sizes and regional focuses [15, 17]. For example, the CDC determined that the autism prevalence rate in the USA is about 1 in 68 children based on only 11 communities [1]. Additionally, most autism prevalence studies do not include undiagnosed individuals [1, 2, 15, 17]. This means that individuals without access to diagnostic centers for socioeconomic or geographic reasons are not reported, resulting in underrepresented statistics. Since location at the city level is considered personally identifiable information, researchers are generally unable to share data with locations attached, which ultimately precludes greater autism epidemiological understanding and accuracy.

Geographic disconnect

Understanding the geographic distribution of autism resources is as difficult as understanding resource gaps. Although many ASD resource directories exist, most are for very small regions (at the city or state level) and can have data integrity drawbacks, including a lack of updated information, incomplete information, and missing resources. More importantly, very few ASD resource directories include critical information pertaining to diagnostic capability. The National Autistic Society United Kingdom's Autism Services Directory [18] is an example of an online resource directory that is autism-specific, relatively comprehensive, and includes key diagnostic information. Replicating such a registry in the USA would not only help complete our understanding of resource distribution, but it would also enable families and individuals with autism to quickly find the best resources near them.

Despite the difficulty, it is still worth finding closer approximations of the geographic distribution of autism and autism resources. Analyses conducted with 47,622 individuals with autism, based on information gathered from online public profiles and social media accounts, and 840 developmental medical centers in the USA, collected through Autism Speaks [13] and Autism Source [14], suggest that resource discrepancies may be much worse than initially thought and that the paucity of resources in various economic communities likely contributes to inequities in a family's ability to access appropriate and necessary therapies, services, and support [19]. The average distance from an individual with a diagnosis of autism to a diagnostic center was estimated at 32 km, and an astonishing 70% of individuals lived no closer than 30 km of a diagnostic center. Assuming geographic variations in autism prevalence rates are relatively modest, it is possible that a majority of individuals with risk for an autism diagnosis live prohibitively far from a diagnostic center––especially with the uneven allocation of 840 diagnostic centers for a nation of 9.85 million squared kilometers [20]. Most likely, there is a large disconnect between resources and individuals with autism that need an official diagnosis and healthcare services.

Mobile solution

To complement the resource lists from Autism Speaks and Autism Source, we have devised a tool, GapMap (http://gapmap.stanford.edu), to obtain more accurate and widespread estimates of geographic variations in autism prevalence rates and resource availability. GapMap is a mobile-first website or an application that renders well and is fully usable from a mobile or tablet device but can also be accessed through a traditional computer. Minorities, households with an income of less than $50,000, and the non-college educated are more likely to use mobile Internet as their primary or only device for Internet access [21]. Individuals in rural areas are less likely to access the Internet, with or without cell phones; however, usage is high enough to warrant developing health-related Internet and mobile applications [22–24]. As such, data collected through GapMap will still be able to reduce bias in prevalence data.

GapMap features a map with overlays of real-time autism prevalence and resource markers. Dynamic features allow visitors to electronically consent, contribute data, find local resources, and learn more about the study. Current estimates of autism prevalence rates have been used to simulate data for the map. Similarly, GapMap's resource bank already contains extracted data from both regional and national pre-existing online resource directories (including Autism Speaks and Autism Source) [13, 14]. This dataset has been further refined by algorithmic categorization, classification (as a center, specialist, or online resource), and deduplication. See Fig. 1 for GapMap's interface.
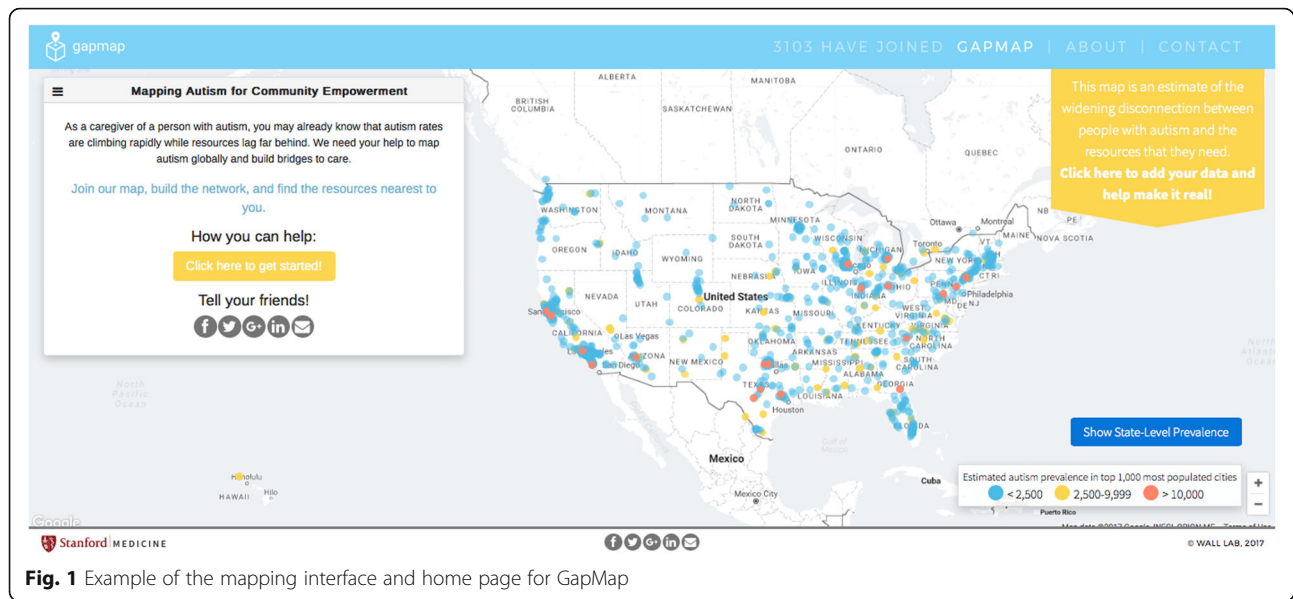
Daniels *et al. Molecular Autism* (2017) 8:55

Page 3 of 6

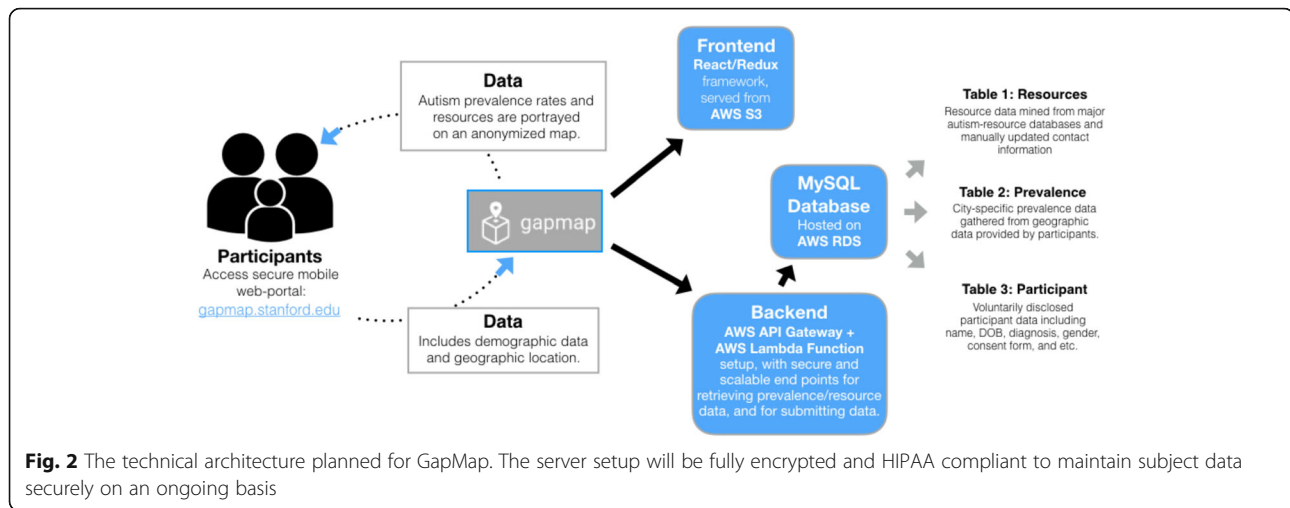**Fig. 1** Example of the mapping interface and home page for GapMap

Neither the prevalence data nor the resource banks are complete, but a simple form lets individuals with autism (or caregivers of a child with autism) submit data. These data include gender, date of birth, location (city and state), specific diagnosis/co-morbid conditions, contact information, and local services that have been used. IP addresses, date and time of submission, and similarity of data submitted will be used to detect duplicate or flag anomalies as potentially falsified data. Participants also provide answers to a machine learning behavioral classification system, which has been shown to match clinical diagnostic outcomes with high frequency [25–29]. Crowdsourced data has been shown to match the quality of expert-curated data with proper instructions for data submission and reasonable validation on input data [30–32].

Local services include medical specialists, therapists, support resources, and "autism-friendly" generic services. After submission, locational data will be anonymously incorporated into the prevalence map; all other data will be securely managed and used to better understand autism resource deficits. In the future, site visitors will also be able to easily add to or edit the autism resource bank and fill in ASD-specific information such as diagnostic capability, target age, and accommodated disorders/disabilities. While resource directories are often difficult and costly to maintain, as new services open and others shut down, crowdsourcing offers a lower cost solution: leverage the collective knowledge of individuals providing, using, and seeking resources. In particular, parents of young children are more likely to search for and share information online [33], and resource providers have an incentive to list themselves for discoverability. IP addresses, submitter account information, and

contribution activity will be tracked and used to detect malicious users, unusual resource deletions or additions, or questionable resource review submissions. Although there may be an incentive for providers to supply "fake reviews" of their services or their competitors, it is a common crowdsourcing problem with existing spam detection and filtering algorithms [34]. This filtering will ensure that questionable resources are removed. In addition, we will validate any organic resources through a machine learning algorithm that will confirm the contact data that was provided by our users that corresponds to what is publicly available on the Internet. If we do not find a resource-match, we will not include the resource in GapMap's database. Our hope is that empowering families and individuals to contribute data allows for a robust and constantly updated global database of autism resources and prevalence rates.

System architecture

Data are encrypted and stored on secure MySQL databases behind a firewall. GapMap is written in React.js and runs on Amazon Web Services Simple Storage Serve (AWS S3). The backend server runs on AWS application program interface (API) Gateway and AWS Lambda. AWS API Gateway executes specific JavaScript packages, novel code that interacts with our SQL database, on AWS Lambda. The MySQL relational database is hosted on Amazon Relational Database Service (RDS) and consists of two main tables. Table 1 holds the resource data, including name/type of resource, geographic coordinates, address, and contact information. Table 2 holds participant data to include specific diagnosis, consent form, geographic location (zip code), and

Daniels *et al. Molecular Autism* (2017) 8:55

Page 4 of 6

Fig. 2 The technical architecture planned for GapMap. The server setup will be fully encrypted and HIPAA compliant to maintain subject data securely on an ongoing basis

other personally identifiable information. See Fig. 2 for an overview of the planned system architecture.

Conclusions

If successful, GapMap could help the families of 3.5 million individuals in the USA [35] quickly and stresslessly find services that range from diagnostic evaluations to a variety of forms of therapy. It could help thousands of individuals recently diagnosed with ASD and their families each year find options for therapy, schooling, insurance, and support. It can also help the approximately 27% of children who remain undiagnosed with ASD by age 8 [36] find the best resource options as they progress through school (e.g., from speech therapy to behavioral therapy to cognitive therapy) and also aid 52 million individuals with ASD worldwide [15] find employment, relationships, and independent living support. In addition, GapMap would allow policy experts to advocate in favor of efforts to increase the number of autism-related clinical practices in resource-poor areas and provide large-scale data for researchers studying the causes of and potential therapies for ASD.

Limitations

There is potential for bias in reporting and data capture that could skew the data incorporated into GapMap over time. To safeguard against biased reporting of autism diagnoses, we have incorporated a set of behavioral questions validated for autism detection by a series of machine learning experiments [26, 28]. We also provide the option to upload a video of the child with autism to enable secondary evaluation of the autism diagnosis using a separate machine learning approach [25, 27, 29]. These approaches maintain a balance between reporting complexity and information content to help ensure high participation while retaining accuracy.

To safeguard against incorporation of inaccurate resource information, we will continue to enhance the GapMap software system to (a) automatically check for the functionality of URLs, emails, and phone numbers, and (b) incorporate and prioritize resources that have well-established reputations. For the latter category, we will use Search Engine Optimization (SEO) rankings to flag Web accessible resource listings that have higher numbers of independent links on external sites and prioritize sites associated with highly regarded academic medical centers, as well as sites with clinical researchers who have PubMed indexed and salient publications in the field of autism care. In addition, as the GapMap population grows, we will enable participants to rate sites listed within their geographic area, providing a "Yelp-like" means to crowd-source the process of vetting listed resources. Over time, this will not only include star ratings but also more detailed information like the average reported waiting time to receive services.

Despite these limitations, utilizing crowdsourced data for the prevalence of diagnosed cases of autism and autism-related resources can impact positively on the quality of and access to robust information, through the involvement of an active and increasingly large population [37–39]. GapMap will be an invaluable tool to members of the autism community, as it can inexpensively and feasibly amass valuable information. With this tool and others like it, we will be able to quantify the geographic disconnect that exists worldwide and leverage this information to innovate targeted strategies that give families answers and ability to act faster and with greater frequency.

Daniels *et al. Molecular Autism* (2017) 8:55

Page 5 of 6

Availability of data and materials
Autism Speaks resource database: https://www.autismspeaks.org//family-services/resource-guide.
Autism Source resource database: http://www.autismsource.org.

Authors' contributions
JD, JS, and NA prepared the initial manuscript. NA and MD created GapMap. DW conceived the project and served as project director. All authors read and approved the final manuscript.

Authors' information
Not applicable.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]Department of Pediatrics, Division of Systems Medicine, Stanford University, Stanford, CA, USA. [2]Department of Biomedical Data Science, Stanford University, Stanford, CA, USA. [3]Department of Computer Science, Princeton University, Princeton, NJ, USA. [4]1265 Welch Rd X141, Stanford, CA 94305, USA.

References

1. Christensen DL, et al. Prevalence and characteristics of autism spectrum disorder among children aged 8 years––autism and developmental disabilities monitoring network, 11 sites, United States, 2012. MMWR Surveill Summ. 2016;65(3):1–23.
2. Christensen DL, et al. Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network. J Dev Behav Pediatr. 2016;37(1):1–8.
3. Russell G, Ford T, Steer C, Golding J. Identification of children with the same level of impairment as children on the autistic spectrum, and analysis of their service use. J Child Psychol Psychiatry. 2010;51:643–51.
4. Wiggins LD, Baio J, Rice C. Examination of the time between first evaluation and first autism spectrum diagnosis in a population-based sample. J Dev Behav Pediatr. 2006;27(2 Suppl):S79–87.
5. Hutton AM, Caron SL. Experiences of families with children with autism in rural New England. Focus Autism Devel Disabil. 2005;20(3):180–9.
6. Mandell DS, Novak MM, Zubritsky CD. Factors associated with age of diagnosis among children with autism spectrum disorders. Pediatrics. 2005; 116(6):1480–6.
7. Bernier R, Mao A, Yen J. Psychopathology, families, and culture: autism. Child Adolesc Psychiatr Clin N Am. 2010;19(4):855–67.
8. Dawson G, Bernier R. A quarter century of progress on the early detection and treatment of autism spectrum disorder. Development and psychopathology. 2013;25(4pt2):1455–1472.
9. Dawson G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. Dev Psychopathol. 2008;20(03):775–803.
10. Dawson G, Burner K. Behavioral interventions in children and adolescents with autism spectrum disorder: a review of recent findings. Curr Opin Pediatr. 2011;23(6):616–20.
11. Rogers SJ, Dawson G. Early start Denver model for young children with autism: promoting language, learning, and engagement. New York: Guilford Press; 2010.
12. Rogers SJ, Dawson G, Vismara LA. An early start for your child with autism: using everyday activities to help kids connect, communicate, and learn. New York: Guilford Press; 2012.
13. Autism Speaks. 2016. http://www.autismspeaks.org. Accessed 29 Nov 2016.
14. Autism Source. 2004. http://www.autismsource.org. Accessed 29 Nov 2016.
15. Baxter AJ, Brugha TS, Erskine HE, Scheurer RW, Vos T, Scott JG. The epidemiology and global burden of autism spectrum disorders. Psychol Med. 2015;45:601–13.
16. Rzhetsky A, Bagley SC, Wang K, Lyttle CS, Cook EH Jr, Altman RB, Gibbons RD. Environmental and state-level regulatory factors affect the incidence of autism and intellectual disability. PLoS Comput Biol. 2014; 10(3):e1003518.
17. Ramsey E, Kelly-Vance L, Allen JA, et al. Autism spectrum disorder prevalence rates in the United States: methodologies, challenges, and implications for individual states. J Dev Phys Disabil. 2016;28:803–20.
18. National Autistic Society: National Autistic Society United Kingdom's Autism Services Directory. 2016. http://www.autism.org.uk/services/helplines/directory.aspx. Accessed 29 Nov 2016.
19. Durkin MS, Elsabbagh M, Barbaro J, Gladstone M, Happe F, Hoekstra RA, Lee LC, Rattazzi A, Stapel-Wax J, Stone WL, Tager-Flusberg H. Autism screening and diagnosis in low resource settings: challenges and opportunities to enhance research and services worldwide. Autism Res. 2015;8(5):473–6.
20. The World Bank. http://data.worldbank.org/indicator/AG.LND.TOTL.K2 (2016). Accessed 30 Nov 2016.
21. Duggan M, Smith A. Cell internet use 2013. Washington, DC: PewResearchCenter; 2013.
22. Greenberg, Alexandra J., et al. Differences in access to and use of electronic personal health information between rural and urban residents in the United States. J Rural Health. 2016.
23. Modipane MB, et al. Technology use among patients in a nonurban southern US HIV clinic in 2015. Telemed E Health. 2016;22(11):965–8.
24. Goh, Jie-Mein, Guodong (Gordon) Gao, and Ritu Agarwal. The creation of social value: can an online health community reduce rural-urban health disparities?. MIS Q 40.1 (2016): 247-263.
25. Fusaro VA, Daniels J, Duda M, DeLuca TF, D'Angelo O, Tamburello J, Maniscalco J, Wall DP. The potential of accelerating early detection of autism through content analysis of YouTube videos. PLoS One. 2014;9(4):e93533.
26. Wall DP, Dally R, Luyster R, Jung JY, DeLuca TF. Use of artificial intelligence to shorten the behavioral diagnosis of autism. PLoS One. 2012;7(8):e43855.
27. Wall DP, Kosmicki J, Deluca TF, Harstad E, Fusaro VA. Use of machine learning to shorten observation-based screening and diagnosis of autism. Transl Psychiatry. 2012;2(4):e100.
28. Duda M, Daniels J, Wall DP. Clinical evaluation of a novel and mobile autism risk assessment. J Autism Dev Disord. 2016;46(6):1953–61.
29. Duda M, Kosmicki JA, Wall DP. Testing the accuracy of an observation-based classifier for rapid detection of autism risk. Transl Psychiatry. 2014;4(8):e424.
30. Comber A, Brunsdon C, See L, Fritz S, McCallum I. Comparing expert and non-expert conceptualisations of the land: an analysis of crowdsourced land cover data. In International Conference on Spatial Information Theory. Scarborough: Springer, Cham; 2013. pp. 243-260.
31. Behrend TS, Sharek DJ, Meade AW, Wiebe EN. The viability of crowdsourcing for survey research. Behav Res Methods. 2011;43(3):800–13.
32. Swan M. Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. J Med Internet Res. 2012;14(2):e46.
33. Stern MJ, Cotten SR, Drentea P. The separate spheres of online health: gender, parenting, and online health information searching in the information age. J Fam Issues. 2011;33(10):1324–50.
34. Mukherjee A, Venkataraman V, Liu B, Glance NS. What yelp fake review filter might be doing? Chicago: InICWSM; 2013.

Daniels *et al. Molecular Autism* (2017) 8:55

Page 6 of 6

35. Buescher AV, Cidav Z, Knapp M, Mandell DS. Costs of autism spectrum disorders in the United Kingdom and the United States. JAMA Pediatr. 2014; 168:721–8.
36. Shattuck PT, et al. Timing of identification among children with an autism spectrum disorder: findings from a population-based surveillance study. J Am Acad Child Adolesc Psychiatry. 2009;48(5):474–83.
37. Jost CC, Mariner JC, Roeder PL, Sawitri E, Macgregor-Skinner GJ. Participatory epidemiology in disease surveillance and research. Rev Sci Tech. 2007;26(3):537. doi: 10.1371/journal.pmed.1000376.
38. Freifeld CC, Chunara R, Mekaru SR, Chan EH, Kass-Hout T, et al. Participatory epidemiology: use of mobile phones for community-based health reporting. PLoS Med. 2010:7(12): e1000376. doi: 10.1371/journal.pmed.1000376.
39. Paolotti D, Carnahan A, Colizza V, Eames K, Edmunds J, Gomes G, Koppeschaar C, Rehn M, Smallenburg R, Turbelin C, Noort S. Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience. Clin Microbiol Infect 2014;20(1):17-21. doi: 10.1111/1469-0691.12477.