

ZikaHack 2016: A digital disease detection competition

Dillon C Adam¹, Jitendra Jonnagaddala², Daniel Han-Chen³, Sean Batongbacal⁴, Luan Almeida⁵, Jing Z Zhu⁶, Jenny J Yang⁷, Jumail M Mundekkat⁴, Steven Badman⁸, Abrar A Chughtai¹, C Raina MacIntyre¹

¹Integrated Systems for Epidemic Response, UNSW Sydney, Australia

²School of Public Health, UNSW Sydney, Australia

³School of Mathematics & Statistics, UNSW Sydney, Australia

⁴School of Computer Science & Engineering, UNSW Sydney, Australia

⁵Computer Science Department, Federal University of Alagoas, Brazil

⁶Concord Repatriation General Hospital, Sydney, Australia

⁷Faculty of Medicine, UNSW Sydney, Australia

⁸Kirby Institute, UNSW Sydney, Australia

Abstract

Effective response to infectious diseases outbreaks relies on the rapid and early detection of those outbreaks. Invalidated, yet timely and openly available digital information can be used for the early detection of outbreaks. Public health surveillance authorities can exploit these early warnings to plan and co-ordinate rapid surveillance and emergency response programs. In 2016, a digital disease detection competition named ZikaHack was launched. The objective of the competition was for multidisciplinary teams to design, develop and demonstrate innovative digital disease detection solutions to retrospectively detect the 2015-16 Brazilian Zika virus outbreak earlier than traditional surveillance methods. In this paper, an overview of the ZikaHack competition is provided. The challenges and lessons learned in organizing this competition are also discussed for use by other researchers interested in organizing similar competitions.

1 Introduction

Rapid detection of disease outbreaks through disease surveillance is critical for early and effective

prevention and control of potential epidemics. Traditional communicable disease surveillance typically includes elements of case detection, validation, and dissemination, to accurately detect outbreaks and inform subsequent control measures where necessary. These methods however, while highly specific, rely on clinical diagnoses and/or laboratory confirmation, and involve a chain of processing from health providers to government health authorities. This can be a time-consuming process, which while suitable for accurate surveillance of long term trends, is not timely enough for rapid outbreak detection. More timely methods of detecting outbreaks could reduce the delay of disease control measures, particularly important during the early hours and days of an outbreak.

The extraordinary increase in public domain data driven by the accessibility of the internet, smart phones and social media, mean that a wealth of information about our individual and collective lives is more readily available than ever before. Utilizing these data for disease surveillance is referred to as digital disease detection. It has shown increasing promise in the early detection and identification of outbreaks, despite concerns regarding the validity and accuracy of disease predictions (Bernardo et al., 2013; Chunara et al., 2012). Despite these limitations, unofficial public data sources, such as online search engines and social media can provide timely information for the early detection of disease outbreaks and can be used to support the early initiation of traditional surveillance activities (Bernardo et al., 2013; Ginsberg et

al., 2009; Salathe et al., 2013). The World Health Organization (WHO) reports that more than 60% of their initial outbreak reports come from unofficial sources (Christaki, 2015; World Health Organization, 2016).

1.1 Zika & the 2015 Brazilian outbreak

Zika virus (ZIKV) was first identified in monkeys in 1947, and then in humans in 1952 in Uganda (Kirya, 1977). Only occasionally reported in equatorial regions of Africa and Asia (Ioos et al., 2014), the virus was not considered a pathogen of significant public health concern until the 2015 epidemic in Brazil, concurrent with reports of associations with birth defects such as microcephaly. The first signs of a potential ZIKV outbreak occurred in May 2015, as the Pan American Health Organization (PAHO) released an alert for possible ZIKV infection in Brazil following a cluster of non-specific rash in February 2015. The first cluster of microcephaly cases was reported in August 2015 but the association with ZIKV infection wasn't noted until November (Schuler-Faccini et al., 2016). By February 2016, the World Health Organization (WHO) declared the ZIKV epidemic to be a Public Health Emergency of International Concern (World Health Organization (WHO), 2016). It is estimated that anywhere between 10–80% of the Brazilian population (207 million) may have been exposed to ZIKV during the outbreak (Jaenisch et al., 2016; Johansson et al., 2016; Nishiura et al., 2016). If symptoms do appear, they are mostly mild and non-specific such as fever, rash, and joint pain typical of many other arboviral diseases such as dengue and chikungunya (Duffy et al., 2009; Grard et al., 2014; Olson et al., 1981). As such, many women were not aware they had become infected until their baby was born with birth defects nine months later. The risk of microcephaly due to ZIKV is estimated to be low, ranging from 0-5%, but may be as high as 30% (Jaenisch et al., 2016; Johansson et al., 2016; Nishiura et al., 2016; Oliveira Melo et al., 2016; Ventura et al., 2016).

2 ZikaHack 2016

Hackathons are typically short-term competitions which bring together multidisciplinary teams to develop innovative solutions to a defined problem.

Hackathon style competitions have proven effective in enhancing student-centred learning and fostering inter professional development (Kienzler & Fontanesi, 2017; Youm & Wiechmann, 2015). Inspired by this, a digital disease detection competition called 'ZikaHack 2016' was organised and sponsored by The National Health and Medical Research Council's (NHMRC) Centre for Research Excellence in Integrated Systems for Epidemic Response (ISER)¹. Based at University of New South Wales (UNSW) Sydney, ISER conducts applied systems research to enhance collaboration and build capacity in health systems for epidemic response and control. Open to university students world-wide, the competition challenged multidisciplinary student teams to design and develop digital disease detection solutions to detect early signals for the ZIKV outbreak in Brazil earlier than traditionally surveillance methods did, using only publicly available data sources. One of the main tasks for the teams was to identify the Brazilian ZIKV outbreak and formulate the scope for potential early signals.

3 Competition structure

3.1 Overview

The competition was split into two phases: a shortlisting phase, and a development phase. Phase one of the competition was launched in August 2016. Eligible teams of three to six students were tasked to submit an application including a proposal of no more than 3,000 words describing their solution to detect an early ZIKV surveillance signal. A key requirement for entry was the cross-disciplinary background and mixed study level of the student teams: team composition had to include both undergraduate and postgraduate students with at least one student from science, technology, engineering or math (STEM), and another from a health-related program. Students had to be enrolled at a recognised university within their country at the time of application. Only applications in English were accepted. For exact eligibility criteria used, please refer to Appendix A.

Details of the competition such as conditions of entry and proposal requirements were posted on the

¹ <https://sphcm.med.unsw.edu.au/centres-units/centre-research-excellence-epidemic-response>

ISER website. The competition was widely publicized on the university website and emails were sent to colleagues working at other universities in Australia and overseas for promotion. A local company, thinkable.org, was also contracted to promote and field applicants for the competition.

3.1 Phase One

Phase one of the competition was launched in August 2016. Students were given approximately five months to form teams and submit a proposal before the closing date of 30 November 2016. Applications were reviewed for eligibility and ranked in a blinded process by a panel of four judges independently. Criteria for the ranking was general: a demonstrated understanding of the ZIKV problem in Brazil, the contest brief and originality of the proposed solution. A median-rank score for each submission was calculated blindly and submissions re-ranked. Judges deliberated over the top performing submissions and selected two finalists to move forward into phase two. Teams were notified on 19 December 2016. A summary of phase one proposals is presented in Table 1.

3.2 Phase Two

Teams shortlisted for phase two were tasked to begin development and implementation of their proposed digital disease detection solutions. The primary criteria for evaluation was the ability of the solution to identify the ZIKV outbreak early compared to the official ZIKV epidemic alert by the WHO in February 2016. Complete solution documentation and source code were also required to be compiled for evaluation. A joint teleconference with the finalists and the competition organizers was held to provide teams with an opportunity to ask questions and discuss any challenges faced. Following the ZIKV outbreak. Teams routinely proposed using machine learning and natural language processing algorithms to develop models for ZIKV prediction from social media sources. However, the queries used and data sources specified differed between teams as did the methods. At least one social media platform was chosen by most teams as a potential data source for digital disease detection however in some cases multiple sources were specified. The

nal submissions were due 30 April 2017. The finalists presented an overview of their solutions and early signals of ZIKV outbreak in Brazil before the same four person judging panel. Presentations were scored according to a final criterion including the ability of the tool to produce an early signal of the ZIKV outbreak earlier than traditional surveillance, originality, feasibility and adaptability of the design. The winning team was notified in early May 2017².

4 Competition results

4.1 Participants

A total of eight proposals were received in phase one. One team was disqualified as they did not meet the entry criteria. Of the seven qualifying teams, the average team size was five persons with 34 total individual student applicants. The large majority student applicants were enrolled at an Australian university (n=32; 94%) and male (n=26; 77.5%). Four out of seven teams collaborated between at least two universities with the remaining three team's composition exclusively of students from the same university. The most common level of study was postgraduate (n=22; 65%) and of those, most were enrolled in a Master level program (n=16; 73%) followed by a doctoral level program (n=6; 27%). Thirty-five percent of applicants were (n=12) enrolled in an undergraduate bachelor program. Across all program levels, public health and computer science were the most common fields of study (n=6; 18% each) followed by data science, information technologies and engineering (n=5; 15% each).

4.2 Proposed solutions

Most but not all teams correctly understood the challenge of identifying early signals of Brazilian most common source was Twitter (n=5/7; 71%), followed by Facebook (n=3/7; 42%). The 'REST' and 'Streaming Twitter' API was commonly proposed (even though it is not possible to obtain the data required for this competition retrospectively) as a means for data extraction. Google trends was another source of data proposed (n=3/7; 47%). Some common query-terms proposed across social

² <https://newsroom.unsw.edu.au/news/health/students-find-early-signals-zika-virus-outbreak>

Team	Data Sources	Algorithms	Language
1	Facebook, Google search, National weather data, Census data	Unnamed machine learning algorithms, Delay differential, Proportional hazards models	Not specified
2	Twitter, PubMed, Google scholar, News sources, Worldclim.org	Unnamed mosquito model, SIR transmission model	Python
3	Facebook, Twitter, Google trends, Google maps, National weather statistics, Mass social events calendar, Wikipedia, HealthMap, CDC	Deep neural networks, Unnamed machine learning algorithms.	Python, R, Matlab
4	United Nations world tourism organization, Government reports	Unnamed clustering algorithm, Random forests, Decision trees.	Python, R
5	Twitter, Global climate data, Google trends	Bayesian Markov network model, Auto regression exogenous model, SVM regression model, Naive Bayes	Not specified
6	Twitter	Deep neural networks	Python
7	Reddit, Twitter, Wikipedia, Instagram	Deep neural networks, Unnamed aberration detection models	Python

Table 1: Summary of phase one proposals

media platforms and search engines included: fever, rash, headache and conjunctivitis. Python was mentioned in all proposed solutions that specified at least one programming language, followed by R and Matlab.

Proposed solutions for ZIKV early detection varied and are summarised in Table 1. Many teams included climate variables into their models as a predictor of mosquito biting risk. For example, Team 1 proposed using an algorithm to predict ideal conditions for a ZIKV outbreak based on historic climate data (mosquito risk) and calculating outliers of a regression analysis using filtered search results. Another incorporated the epidemiological concept of R_0 (expected number of secondary infection produced by primary infection) by calculating the predicted number of ZIKV infections by location using a trained machine learning model from twitter data, and then statistically compared this value to other models with similar symptoms such as dengue and chikungunya. Some solutions however demonstrated a collection of ideas that lacked detail or were inappropriate for desired challenge outcome. Two of the proposed solutions were selected for phase two development.

4.3 Winning Solution: Gadyan

Gadyan (in Australian aboriginal language means "Sydney shellfish") was able to generate early signals, approximately three months before the WHO official alert in early 2016. Gadyan employed a multi-stage pipeline based approach with various components and sub-components incorporated into the solution. Gadyan specifically focused on microcephaly syndromic surveillance and used retrospectively collected relevant data from Google trends, Twitter and Wikipedia for the period Jan 2013 to December 2016. The data extracted from these sources varied in type and formats. Unstructured tweets were extracted from twitter users in Brazil and structured data from the Google trends and Wikipedia. As part of this solution, automatic translation was also performed on Portuguese and Spanish data. The data extracted from these various heterogeneous data sources were appropriately represented in standardized time series (weekly intervals and monthly) formats. The standardized data was further used to generate outbreak alerts. Initially, the alerts were generated using a single data source and conventional aberration detection algorithms. However, the better early warning alerts were possible by combining all the data sources

and using change point detection algorithms instead of standard aberration detection algorithms. The official microcephaly surveillance data from WHO/PAHO was used to perform correlation analysis and assessment of the outbreak alerts during the development of the solution. The developed Gadyan solution is subjected to various limitations but can be extended to detect other disease outbreaks.

4.4 Runner Up

The second finalist and runner-up developed an Outbreak Confidence Distribution (OCD) model to compute the likelihood of a ZIKV outbreak by location (Brazilian states). The solution used a combination of retrospective google search term queries downloaded in both Spanish and Portuguese to train a stacked Machine Learning model with Random Forests and Neural Nets. The model was trained on data sourced from three Brazilian states and then tested on the remaining states. The solution produced signals of possible ZIKV outbreak as early as September 2014 in the state of Espírito Santo. The early signal of ZIKV varied for each state, but the model commonly identified November 2014 as the potential origin of the outbreak in Brazil. The solution however suffered from some noisy signals.

5 Discussion

The results demonstrate that student-centred multidisciplinary teams can provide unique and innovative solutions to challenging digital disease detection problems. There are many advantages in carrying out competitions such as this in an academic setting. Hackathons are frequently reported as hubs for student research innovations (Artiles & Wallace, 2013; Briscoe, 2014) and the results of ZikaHack concur that university organised competitions with prize monies can incentivise talented students to apply for university competitions. Overall, the number of unique solutions proposed satisfied the goals of competition organisers.

Twitter as the choice source of public domain data was unsurprising. Many past efforts to create digital disease detection (DDD) tools have used Twitter as a primary data source to track epidemics (Aramaki et al., 2011; Jonnagaddala et al., 2016;

Lamos et al., 2010). Despite its popularity, the effectiveness of twitter for DDD is questionable. For users, only 1% of the data can be publicly accessed (0.2% of which is geocoded), query terms used to retrieve data can lead to bias, and most tweets originate from the United States (Al-garadi et al., 2016; Romano et al., 2016). In addition, Twitter is also a very difficult source to perform microcephaly syndromic surveillance because identification of pregnant women on social media is a challenging task (Huang et al., 2017). Google search trends can also suffer from similar biases. However, many proposals attempted to overcome issues of bias by integrating additional data such as climate into their models or weighting certain parameters and data points over others. For example, tweets that referenced rash and conjunctivitis would be weighted more than headache or haemorrhage due to their greater association with ZIKV infection compared with similar syndromes caused by chikungunya virus and dengue. This was recognised correctly as the primary challenge in identifying an early ZIKV signal by most teams.

The popularity of Python as the programming language of choice by teams was also unsurprising. It is well agreed that Python's relatively easy syntax, speed, and vast array of available libraries are particularly suited for DDD. Python as a language for data science and machine learning has recently surpassed R and all signs point to this trend continuing (Granville, 2017). While no restrictions on programming language was specified in the ZikaHack competition documents, if a single language is preferred to potentially ease administration and judging, future organisers should feel comfortable selecting Python.

We observed various issues and challenges in organising the ZikaHack competition. The eligibility requirements may have proved a barrier to some applicants, specifically, the requirement that students need to form a multi-disciplinary team of three to six members. Despite a comfortable application period, some interested students struggled to form teams and contacted the organisers to enquire about potential exemptions. While no exceptions were offered, a Facebook group and event page was organised shortly after launch to help students or single-discipline groups find potential team mates. The idea of using social media platform proved effective and we observed good engagement and

communication between interested students. Those considering organising similar cross-disciplinary competitions would be advised to consider the use of social media to stimulate collaborations prior to launch.

Following the close of phase one and as the competition progressed into phase two, it became clear during that workloads were shared unevenly across teams, and often substantial contributions were from single students. In one instance, 70% of the work was formally declared as the work of a single team member, with as little as 5% coming from another in the same team. In another case, planning, formulation, and implementation of the digital disease detection models was the responsibility of one student, and all other team members declared equal responsibility for reviewing the model implementation and background. This may be evidence of a heavy technical burden required from computer science or engineering students to complete the challenge, and may have reduced the number of proposals submitted, however this cannot be proven. With this in mind, future competitions like ZikaHack with specific cross-disciplinary entry requirements may consider specifying a maximum and minimum declared workload per student to improve workload balance. Alternatively, loosening the multidisciplinary team requirements may improve balance but might also have adverse effects, such as a lack of understanding of the core challenges, in this case specifically related to ZIKV and health. The entry requirements in ZikaHack however were not without purpose: as mentioned, one of the primary goals of ISER (competition organiser) is to enhance collaboration and build capacity in health systems research for epidemic response and control. As such, it was important that there was cross-disciplinary engagement across student faculties, particularly between health and STEM, and to advocate cross-faculty research in the various areas of disease surveillance.

Female student participation was low in ZikaHack. Twenty-four percent (8/34) of all individual student applicants was female. Approximately 18% of current Australian Information Technology, Engineering and related technology students are female (Australian Government Department of Education and Training, 2017). We could therefore expect a similar distribution of gender across entries. While 24% of all applicants were female,

91% were from health-related fields meaning only 9% of female applicants were from computer science and engineering related areas. Studies have shown female participation in hackathons are significantly underrepresented when compared to enrolments (Richard et al., 2015), the reasons for which are not well examined, but may be symptomatic of the broader cultural challenges faced by females in predominately male fields such as reduced confidence, prejudice and stereotyping (Irani, 2004). Others suggest the underlying cause may be similar to societies' wider issues of underrepresentation of female computer science and engineering graduates employed in industry and academia. (Briscoe, 2014) Efforts to increasing the gender diversity in hackathon style competitions may therefore benefit from diversity quotas which may help resolve this imbalance. Alternatively, explicitly encouraging female entrants may reduce feelings of being unwelcome and the perceptions of a 'boys-club' (Warner & Guo, 2017). The reasons for reduced female participation in ZikaHack however is purely speculative, as participants were not interviewed.

6 Conclusion

ZikaHack 2016 was a unique, rewarding and successful adaption of the hackathon format and could be replicated yearly either as a continuation, by building on the work of previous years, or a source for new ideas by varying the specific challenge. As a competition with multi-disciplinary requirements, ZikaHack exposed students to concepts outside their field of study, which in turn may inspire participants into alternative research areas; for example, the proposals submitted here could serve as a platform for a future research proposals. In the case of ZikaHack, while never the original goal, research opportunities were offered to some finalists who exceeded expectations. Academics considering organising similar events might also consider how such competitions, perhaps organised yearly with associated monies, may initiate and sustain their areas of research into the future.

Acknowledgments

Funding for this competition was provided by the National health and medical research council's

Centre for Research Excellence, Integrated Systems for Epidemic Response (ISER), grant number APP1107393.

References

- Al-garadi, Mohammed Ali, Khan, Muhammad Sadiq, Varathan, Kasturi Dewi, Mujtaba, Ghulam, & Al-Kabsi, Abdelkodose M. (2016). Using online social networks to track a pandemic: A systematic review. *Journal of Biomedical Informatics*, 62, 1-11.
- Aramaki, Eiji, Maskawa, Sachiko, & Morita, Mizuki. (2011). *Twitter catches the flu: detecting influenza epidemics using Twitter*. Paper presented at the Proceedings of the conference on empirical methods in natural language processing.
- Artiles, Jessica A, & Wallace, David R. (2013). Borrowing from hackathons: overnight designathons as a template for creative idea hubs in the space of hands-on learning, digital learning, and systems re-thinking. *WEEF, Cartagena*.
- Australian Government Department of Education and Training (2017). uCube: Enrolment Count by Field Of Education by Gender. Retrieved from <http://highereducationstatistics.education.gov.au/>
- Bernardo, T. M., Rajic, A., Young, I., Robiadek, K., Pham, M. T., & Funk, J. A. (2013). Scoping review on search queries and social media for disease surveillance: a chronology of innovation. *Journal of Medical Internet Research*, 15(7), e147.
- Briscoe, Gerard. (2014). Digital innovation: The hackathon phenomenon.
- Christaki, E. (2015). New technologies in predicting, preventing and controlling emerging infectious diseases. *Virulence*, 6(6), 558-565.
- Chunara, R., Andrews, J. R., & Brownstein, J. S. (2012). Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak. *American Journal of Tropical Medicine & Hygiene*, 86(1), 39-45.
- Duffy, Mark R., Chen, Tai-Ho, Hancock, W. Thane, Powers, Ann M., Kool, Jacob L., Lanciotti, Robert S., . . . Hayes, Edward B. (2009). Zika Virus Outbreak on Yap Island, Federated States of Micronesia. *New England Journal of Medicine*, 360(24), 2536-2543.
- Ginsberg, J., Mohebbi, MH., Patel, RS., Brammer, L., Smolinski, MS., & Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, 457, 1012-1015.
- Granville, V. (2017, August 30 2017). Python Overtakes R for Data Science and Machine Learning. Retrieved from <http://www.datasciencecentral.com/profiles/blogs/python-overtakes-r-for-data-science-and-machine-learning>
- Grard, G., Caron, M., Mombo, I. M., Nkoghe, D., Mboui Ondo, S., Jiolle, D., . . . Leroy, E. M. (2014). Zika virus in Gabon (Central Africa)--2007: a new threat from *Aedes albopictus*? *PLoS Negl Trop Dis*, 8(2), e2681.
- Huang, Yi-Jie , Su, Chu Hsien , Chang, Yi-Chun , Ting, Tseng-Hsin , Fu, Tzu-Yuan , Wang, Rou-Min , . . . Hsu, Wen-Lian. (2017). *Incorporating Dependency Trees Improve Identification of Pregnant Women on Social Media Platforms*. Paper presented at the Proceedings of the International Workshop on Digital Disease Detection using Social Media 2017, Taipei, Taiwan.
- Ioos, S., Mallet, H. P., Leparce Goffart, I., Gauthier, V., Cardoso, T., & Herida, M. (2014). Current Zika virus epidemiology and recent epidemics. *Medecine et Maladies Infectieuses*, 44(7), 302-307.
- Irani, Lilly. (2004). *Understanding gender and confidence in CS course culture*. Paper presented at the ACM SIGCSE Bulletin.
- Jaenisch, Thomas, Rosenberger, Kerstin Daniela, Brito, Carlos, Brady, Oliver, Brasil, Patricia, & Marques, Ernesto. (2016). Estimating the risk for microcephaly after Zika virus infection in Brazil. *Bulletin of the World Health Organization*.
- Johansson, M. A., Mier-y-Teran-Romero, L., Reefhuis, J., Gilboa, S. M., & Hills, S. L. (2016). Zika and the Risk of Microcephaly. *New England Journal of Medicine*, 375(1), 1-4.
- Jonnagaddala, Jitendra, Jue, Toni Rose , & Dai, Hong-Jie. (2016). *Binary classification of Twitter posts for adverse drug reactions*. Paper presented at the Proceedings of the Social Media Mining Shared Task Workshop at the Pacific Symposium on Biocomputing, Big Island, HI, USA.
- Kienzler, Hanna, & Fontanesi, Carolyn. (2017). Learning through inquiry: a Global Health Hackathon. *Teaching in Higher Education*, 22(2), 129-142.
- Kirya, B. G. (1977). A yellow fever epizootic in Zika forest, Uganda, during 1972: Part 1: Virus isolation and sentinel monkeys. *Trans R Soc Trop Med Hyg*, 71(3), 254-260.
- Lamos, Vasileios, De Bie, Tijn, & Cristianini, Nello. (2010). Flu detector-tracking epidemics on Twitter. *Machine Learning and Knowledge Discovery in Databases*, 599-602.
- Nishiura, H., Mizumoto, K., Rock, K. S., Yasuda, Y., Kinoshita, R., & Miyamatsu, Y. (2016). A theoretical estimate of the risk of microcephaly during pregnancy with Zika virus infection. *Epidemics*, 15, 66-70.

- Oliveira Melo, A. S., Malinger, G., Ximenes, R., Szejnfeld, P. O., Alves Sampaio, S., & Bispo De Filippis, A. M. (2016). Zika virus intrauterine infection causes fetal brain abnormality and microcephaly: Tip of the iceberg? *Ultrasound in Obstetrics and Gynecology*, 47(1), 6-7.
- Olson, J. G., Ksiazek, T. G., Suhandiman, & Triwibowo. (1981). Zika virus, a cause of fever in Central Java, Indonesia. *Trans R Soc Trop Med Hyg*, 75(3), 389-393.
- Richard, Gabriela T, Kafai, Yasmin B, Adleberg, Barrie, & Telhan, Orkan. (2015). *StitchFest: Diversifying a College Hackathon to broaden participation and perceptions in computing*. Paper presented at the Proceedings of the 46th ACM Technical Symposium on Computer Science Education.
- Romano, Sara, Di Martino, Sergio, Kanhabua, Nattiya, Mazzeo, Antonino, & Nejd, Wolfgang. (2016). *Challenges in detecting epidemic outbreaks from social networks*. Paper presented at the Advanced Information Networking and Applications Workshops (WAINA), 2016 30th International Conference on.
- Salathe, M., Freifeld, C. C., Mekaru, S. R., Tomasulo, A. F., & Brownstein, J. S. (2013). Influenza A (H7N9) and the importance of digital epidemiology. *New England Journal of Medicine*, 369(5), 401-404.
- Schuler-Faccini, L., Ribeiro, E. M., Feitosa, I. M., Horovitz, D. D., Cavalcanti, D. P., Pessoa, A., . . . Brazilian Medical Genetics Society-Zika Embryopathy Task, Force. (2016). Possible Association Between Zika Virus Infection and Microcephaly - Brazil, 2015. *MMWR Morb Mortal Wkly Rep*, 65(3), 59-62.
- Ventura, C. V., Maia, M., Bravo-Filho, V., Góis, A. L., & Belfort, R., Jr. (2016). Zika virus in Brazil and macular atrophy in a child with microcephaly. *The Lancet*, 387(10015), 228.
- Warner, Jeremy, & Guo, Philip J. (2017). *Hack. edu: Examining How College Hackathons Are Perceived By Student Attendees and Non-Attendees*. Paper presented at the Proceedings of the 2017 ACM Conference on International Computing Education Research.
- World Health Organization. (2016). Epidemic intelligence - systematic event detection. . Retrieved from <http://www.who.int/csr/alertresponse/epidemicintelligence/en/>
- World Health Organization (WHO). (2016). WHO statement on the first meeting of the International Health Regulations (2005) (IHR 2005) Emergency Committee on Zika virus and observed increase in neurological disorders and neonatal malformations.

Retrieved from <http://www.who.int/mediacentre/news/statements/2016/1st-emergency-committee-zika/en/>

- Youm, J., & Wiechmann, W. (2015). The Med AppJam: a model for an interprofessional student-centered mHealth app competition. *J Med Syst*, 39(3), 34.

Appendix A. Eligibility Criteria

The following eligibility Criteria was used for the Zik-aHack 2016 competition.

- Student team has 3 to 6 enrolled students (who must all be enrolled at the time of the Phase 1 Submission Date of 30 November 2016)
- There is a single nominated team leader
- Team includes undergraduate and postgraduate students
- Team includes students from the following two discipline areas: STEM (science, technology, engineering, mathematics) and health related (medicine, nursing, public health, allied health) disciplines
- Must be studying at a registered university and recognised within its country as a university.
- Letter of support including verification of the student's status of enrolment using the template provided for each team member is attached to the application.
- Application in English
- No team member had a direct connection with any investigator or affiliate of ISER (such as a student-supervisor relationship)
- The work has been done entirely by the student team, with no other assistance.
- All students agreed to be named as part of the team