Online Journal of Public Health Informatics

Journal Impact Factorâ ¢ (JIF) (2021): 7.08
Volume 15 (2023) ISSN: 1947-2579 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Viewpoint

Framework for Classifying Explainable Artificial Intelligence (XAI) Algorithms in Clinical Medicine

Thomas Gniadek[1*], MD, PhD; Jason Kang[1*], MD; Talent Theparee[1*], MD; Jacob Krive[2,3,4,5*], PhD

[1]Department of Pathology and Laboratory Medicine, NorthShore University Health System, Evanston, IL, United States

[2]Department of Biomedical and Health Information Sciences, University of Illinois at Chicago, Chicago, IL, United States

[3]Department of Health Information Technology, NorthShore University Health System, Evanston, IL, United States

[4]Department of Health Informatics, Dr Kiran C Patel School of Osteopathic Medicine, Nova Southeastern University, Fort Lauderdale, FL, United States

[5]Pritzker School of Medicine, University of Chicago, Chicago, IL, United States

[*]all authors contributed equally

Corresponding Author:
Jacob Krive, PhD
Department of Biomedical and Health Information Sciences
University of Illinois at Chicago
1919 W Taylor St 233 AHSB
MC-530
Chicago, IL, 60612
United States
Phone: 1 312 996 1445
Email: krive@uic.edu

Abstract

Artificial intelligence (AI) applied to medicine offers immense promise, in addition to safety and regulatory concerns. Traditional AI produces a core algorithm result, typically without a measure of statistical confidence or an explanation of its biological-theoretical basis. Efforts are underway to develop explainable AI (XAI) algorithms that not only produce a result but also an explanation to support that result. Here we present a framework for classifying XAI algorithms applied to clinical medicine: An algorithm's clinical scope is defined by whether the core algorithm output leads to observations (eg, tests, imaging, clinical evaluation), interventions (eg, procedures, medications), diagnoses, and prognostication. Explanations are classified by whether they provide empiric statistical information, association with a historical population or populations, or association with an established disease mechanism or mechanisms. XAI implementations can be classified based on whether algorithm training and validation took into account the actions of health care providers in response to the insights and explanations provided or whether training was performed using only the core algorithm output as the end point. Finally, communication modalities used to convey an XAI explanation can be used to classify algorithms and may affect clinical outcomes. This framework can be used when designing, evaluating, and comparing XAI algorithms applied to medicine.

(Online J Public Health Inform 2023;15:e50934) doi:10.2196/50934

KEYWORDS

explainable artificial intelligence; XAI; artificial intelligence; AI; AI medicine; pathology informatics; radiology informatics

Introduction

Algorithmic classifiers like artificial neural networks were first implemented many years ago [1]. Recently, unsupervised neural networks have allowed context-agnostic training and deployment. Without the need to embed a priori knowledge of the real-world system being studied, the use of these applications has expanded rapidly, and there has been much excitement about artificial intelligence (AI) algorithms in nearly every industry, including medicine.

Meanwhile, government policy that incentivizes the use of electronic medical record systems expanded the availability of digital health care information [2]. This created an environment where data analysis, predictive analytics, and ultimately AI can readily influence the interpretation of patient data and potentially prevent errors in real time during the course of clinical care [3]. Along these lines, radiologists, and to a lesser extent pathologists, are increasingly using image analysis algorithms as an assistive technology for image interpretation [4-6]. These technologies, rather than feeding into misconceptions about

threats and capabilities of AI, could potentially put radiologists and pathologists at the forefront of purposeful AI innovation [7].

Initially, AI may seem like a threat to health care jobs, removing providers from the decision-making process by introducing algorithms that function as a "black box" [8]. With this perceived threat are concerns about patient safety, some stemming from comparisons to non–health care applications of AI. Like any system, AI is not infallible. For example, early versions of self-driving automobile algorithms may have caused accidents [9].

The practice of clinical medicine remains an "art" where decisions of licensed providers are relied upon to ensure patient safety. Unfortunately, in contrast to transparent, rule-based systems, a trained AI model is not transparent to a clinician [10]. Therefore, there are currently efforts to find a middle ground that combines human involvement and AI in a complementary manner [11]. For example, AI might be used to generate insights not always or easily identified by a human, but a human would still determine their significance [12,13]. In this way, AI becomes a tool used by a clinician.

Multiple countries have passed or proposed regulations on the use of algorithms in clinical medicine. Under the US Food, Drug, and Cosmetic Act, an algorithm can be classified as a "nonregulated medical device" if it meets certain criteria; otherwise, it may represent a regulated medical device. One of the key criteria is whether the algorithm is "intended for the purpose of enabling such health care professional to independently review the basis for such recommendations that such software presents so that it is not the intent that such health care professional rely primarily on any of such recommendations to make a clinical diagnosis or treatment decision regarding an individual patient" [14]. It remains to be seen how the FDA enforces this criterion on a case-by-case basis, and regulations may change over time. Similarly, the UK Department of Health and Social Care has issued robust guidance for best practices in digital health care innovation [15]. One of the key elements of this guidance is transparency about algorithm limitations, algorithm type, and evidence of effectiveness. Because of these regulatory frameworks, concerns about medical malpractice issues, and the general awareness that algorithm predictions are not always correct, there is a growing recognition that AI algorithms should allow health care providers to independently review some form of explanation of their core results [16].

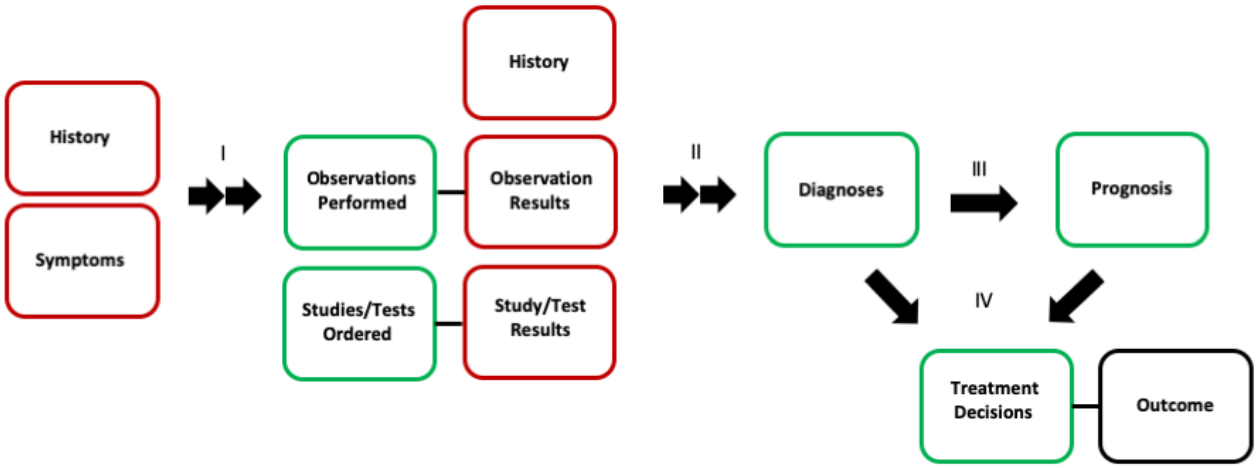
Recently, efforts began to build AI algorithms that allow humans to evaluate the significance of their results, with the goal of better integration and communication between the two. Most notably, the US Defense Advanced Research Projects Agency (DARPA) has called for further development of "explainable artificial intelligence" (XAI) [17]. The core algorithmic result or prediction is provided to the user along with an explanation that is intended to convey insight into the confidence of the core prediction, increase a user's understanding of the real-world process being studied, or both [18].

With its many benefits, XAI also brings added complexity in the form of process-specific outputs and integration with a subject matter expert end user. Not only does this elevate the importance of partnerships between clinicians and AI developers, it also raises the somewhat paradoxical possibility that algorithms with inferior core predictive power may perform better if the explanations provided result in superior outcomes overall. Furthermore, the clinical decision points supported by XAI as well as the manner in which explanations are provided to the user may differ greatly between algorithms and influence their efficacy. Here, we propose a framework for classifying XAI algorithms in clinical medicine in order to simplify this additional complexity and allow for performance evaluation of XAI in clinical practice.

Clinical Scope

The ultimate scope of clinical medicine is to prolong and improve the quality of human life. Within this, there are many decisions and actions that can be evaluated independently (eg, ordering a test, prescribing a medication, performing a surgery). XAI algorithms can be classified based on which step(s) in the clinical care pathway they support (see Figure 1). A single algorithm may provide outputs that encompass multiple areas of clinical scope. Defining clinical scope is critical for XAI, because it will determine which individuals on clinical care teams will be best suited to interact with the algorithm and evaluate the explanations provided. Furthermore, the ultimate impact of XAI on clinical outcomes will be limited by the potential impact of the process steps that an algorithm supports.

Figure 1. Clinical scope for XAI algorithms. XAI algorithms can be classified based on which steps in the clinical decision-making process they support. A simplified process flow map divides clinical decision-making into information (boxes) and information processing (arrows). Information processing steps (I-IV) can involve both human cognitive processing and computerized algorithms. Disease process evolution introduces biologic time dependency (red boxes), leading to a requirement for repeated information processing over time (double arrows). Some recorded information more directly reflects underlying disease (red boxes), while some is mainly the result of information processing (green boxes). Clinical outcome reflects underlying biology, the performance of the entire process, and the effectiveness of treatments. XAI fundamentally influences the information processing steps (I-IV) in partnership with clinicians. XAI performance can be evaluated at each information processing step or studied in the context of overall outcome. Performance of tests and treatments (black lines) are assumed to be static; however, they can be incorporated as inputs into a decision process. XAI: explainable artificial intelligence.



Clinical Insight

Explanations provided by XAI algorithms should aim to provide evidence and ultimately insight to the end user. In the case of pathology, generation of insight to assist clinicians can assist with formation of differential diagnoses, quantitative classification of features, risk prediction, and identification of features imperceptible to the human observer [19]. Both the content of the information and its delivery will determine effectiveness. Evidence can be presented in the form of empiric assessments of statistical confidence, such as a *P* value. Alternatively, an algorithm could provide an assessment of the degree of association between the current patient's data and historical groups of patients or established disease mechanisms (see Table 1).

Clinical providers evaluate empiric assessments of confidence differently than associative power, and the existence of a high degree of uncertainty in any patient-specific medical prediction necessitates a continued role for the "art of medicine" in the form of decision-making by end users. This is due to an incomplete accounting for biological factors that influence disease processes, incomplete documentation of observable factors in the electronic medical record, and the importance of the doctor-patient relationship in clinical care [20]. As a result, associative explanations may be more powerful in certain situations, since an association may support a nonquantifiable opinion held by provider or patient.

Table 1. Classifying explainable artificial intelligence explanations by type. The explanations produced by an explainable artificial intelligence algorithm can provide addition information to a clinician in 3 general ways.

XAI[a] explanation type	XAI explanation output	Primary task for clinician	Benefit to clinician
Empiric	Statistical confidence based on historical sample data	Weigh the degree of confidence provided with risks, benefits, and training data used	Assess the validity of the prediction
Population associative	Association between signs and symptoms of a patient with historical groups of patients	Assess the validity of associating this patient with historical groups of patients	Consider alternative options processed by the algorithm
Mechanism associative	Association with known pathologic mechanism(s)	Assess the validity of the pathologic mechanism(s) and diagnoses proposed	Assess validity of the prediction and consider alternatives using established medical paradigms

[a]XAI: explainable artificial intelligence.

Training and Validation

The loss of context and end user agnostic efficiency of traditional AI algorithms remains a great challenge to the initial design and implementation of XAI. In fact, the meaning of model validation in medicine differs from the traditional validation process typically undertaken in technology fields in that it refers to validation relative to patient outcomes and evidence-based medicine principles—not just whether outcomes are technically correct, match a reference method, or agree with

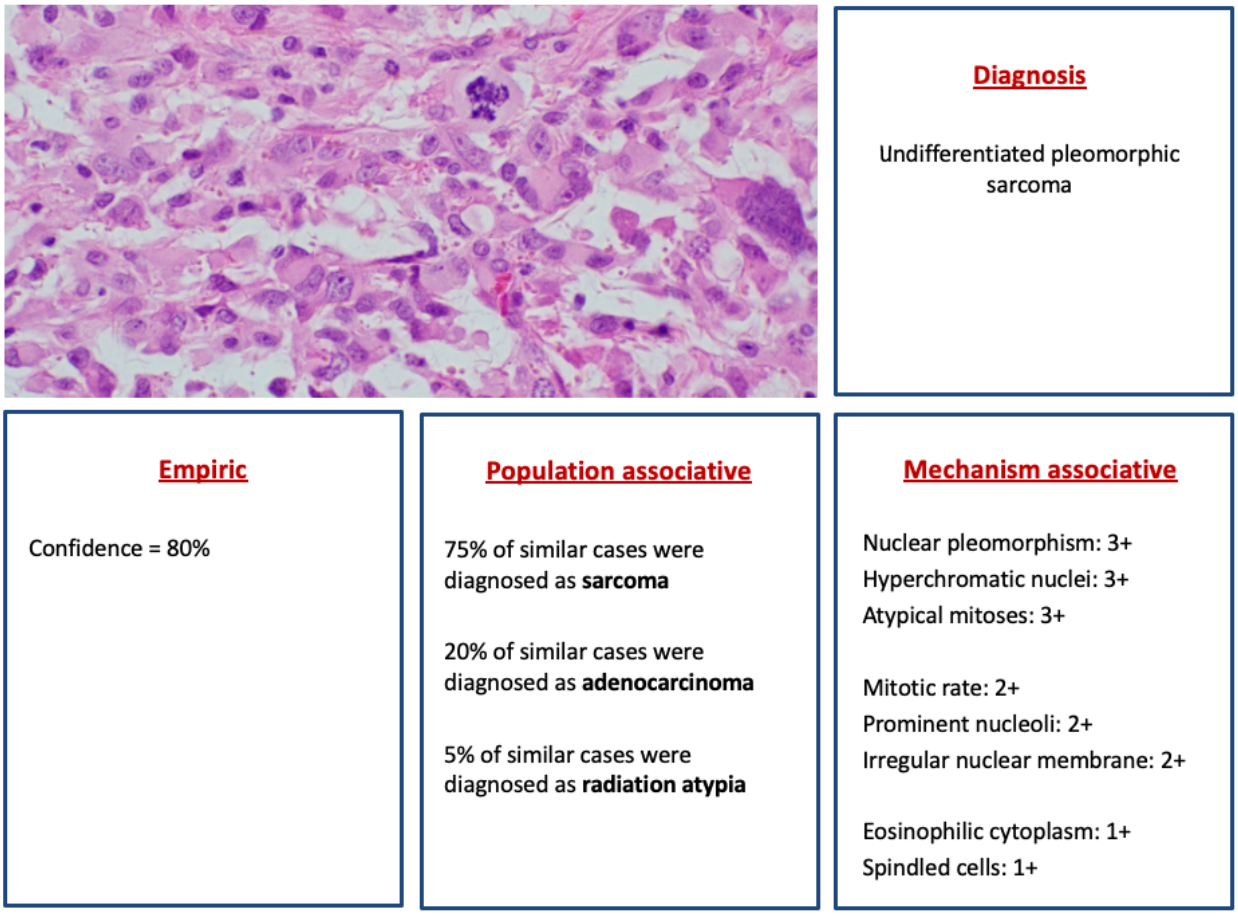
expectations [21]. Ultimately, only patient outcomes can confirm whether the model is valid and whether AI investment is or was a worthy investment. Therefore, XAI takes special meaning in such evidence-based validation processes, since explainable analytics will help support outcomes or facilitate corrections and adjustments. Likely, the development of context-specific XAI will evolve from traditional AI in phases, each supposing a core algorithm output in addition to some form of explanation: phase I will involve traditional AI training and validation; phase II will involve traditional AI training and XAI validation, taking into account end-user actions; and phase III will involve XAI training and validation, both taking into account end-user actions.

During the final phase of XAI development as described above, the algorithm will train not to maximize the predictive power of the core algorithm output but to maximize the outcome of the combined effects of core output, explanation, and end-user actions. It is during this phase of development that XAI implementations may regain some degree of the context-agnostic advantages of traditional AI, since the behavior of the end-user context expert can be studied by the algorithm during validation.

Example 1: Anatomic Pathology

Anatomic pathologists interpret microscopic tissue morphology based on architectural and cytomorphologic criteria shown to correlate with pathologic diagnoses such as cancer. Criteria may include features such as hyperchromatic nuclei, high mitotic rate, and irregular nuclear membrane contours. Unfortunately, none of these features are 100 percent specific for a particular diagnosis like cancer, since nonneoplastic conditions may produce similar cellular features. Additionally, noninvasive premalignant conditions such as carcinoma in situ can contain individual cells that appear morphologically identical to cells within an invasive cancer. Incorporating concepts of XAI into digital anatomic pathology workflows will aid pathologists not only in making the correct diagnosis, but also in considering alternative diagnoses and recognizing potential diagnostic pitfalls (see Figure 2). Potentially, XAI systems can also incorporate ancillary information, such as clinical history, immunohistochemistry staining, and genomic testing, to aid the pathologist.

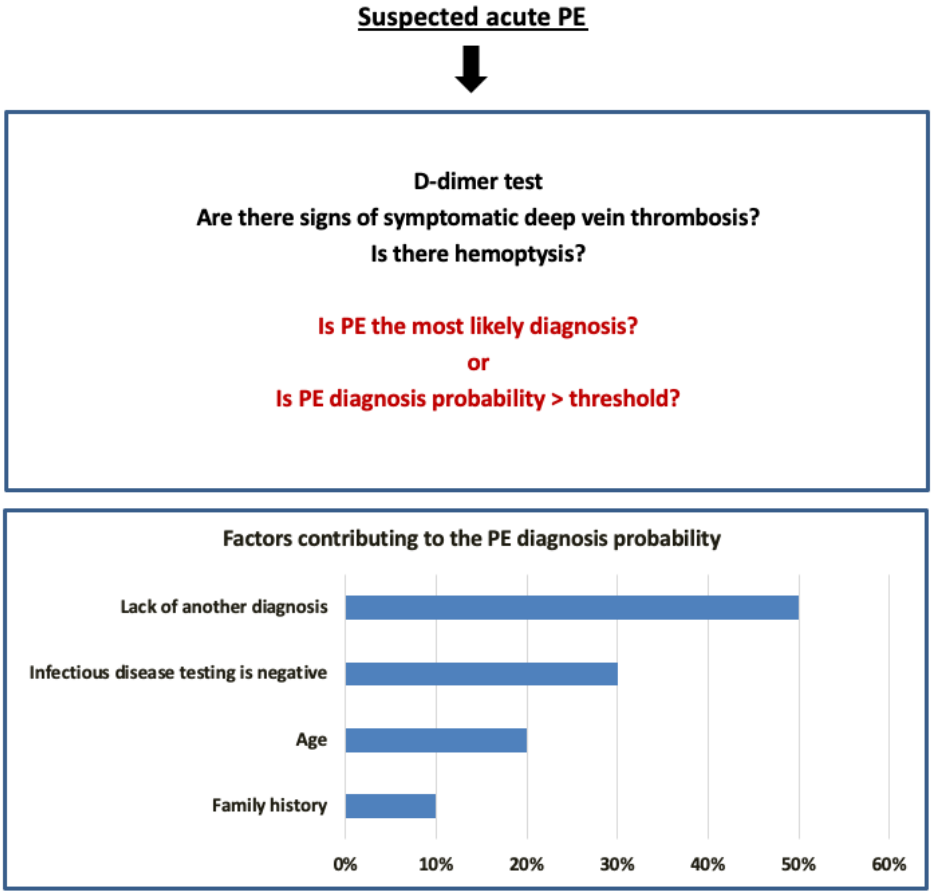
Figure 2. Illustrative example of 3 types of XAI output applied to anatomic pathology. XAI core algorithm output is shown as a diagnosis. Several forms of output explanation are succinctly outlined beneath the image, enabling a physician to make a visual interpretation in conjunction with immediate access to an explanation under multiple categories. “Empiric” information provides overall accuracy expressed as a single number; “population associative” provides a more detailed glimpse into the “black box” result; “diagnosis” relates to other cases an algorithm has access to; “mechanism associative” maps the AI process onto clinically relevant features found in the image (scored based on degree of association, 1 to 3+). XAI: explainable artificial intelligence.



Example 2: Diagnostic Management

One of the most difficult tasks for a clinician is to identify which patients should undergo screening tests and which should not [22]. This is particularly difficult when the condition screened for has a high mortality rate if not recognized, but the screening test is expensive and not without risks. Such a situation exists in deciding whether to screen for pulmonary embolism using computed tomography pulmonary angiography [23]. As a result, algorithms have been developed to aid clinical decision-making, but a clinician's assessment of whether pulmonary embolism is the most likely diagnosis plays a large role in determining a patient's score and management. Scenarios like this represent an opportunity for XAI to contribute toward more accurate assessments of pretest diagnostic likelihood (see Figure 3).

Figure 3. Diagnostic management. Possible modification to the YEARS algorithm for decisions on screening for PE by computed tomography. Rather than relying on clinician assessment of whether PE is the most likely pretest diagnosis, simple scoring algorithms can use an explainable artificial intelligence core algorithm output to assess pretest probability in the context of well-defined historical patient populations. Furthermore, the contribution of factors contributing to the core probability assessment can be displayed. Users can then assess whether each factor is valid, which may influence their assessment of the core algorithm output. For example, factors may be considered invalid if the electronic medical record is recognized as being incomplete or inaccurate. PE: pulmonary embolism.



Conclusions

The 2 recognized advantages of XAI over traditional AI can be summarized as insight into the statistical significance of a core algorithm output and mechanistic insight into the process being studied. It has been suggested that forcing AI to provide mechanistic understanding could decrease the predictive power of the algorithm itself. This may be true in a situation where algorithm inputs include all data relevant to the real-world process; however, clinical medicine remains an area where digitized information is incomplete relative to the totality of factors influencing human disease. Therefore, humans will likely remain the ultimate "trusted" decision-makers during critical, high-risk decisions in clinical care for the foreseeable future. In this framework, even clinical algorithms that are approved as regulated medical devices will remain ancillary to the human practice of medicine. XAI offers the potential to improve not the predictive power of black box algorithms but rather their usefulness as a tool for clinical providers, offering the opportunity to classify and categorize data [24], as well as ensure meaningful feedback that fits clinical workflows [25]. Information should include identification of tasks, the nature and purpose of the tasks, their outcome, and methods applied to produce the outcome [26].

Medical leaders have discussed the need for a "learning health care system" for many years. The development of XAI offers the potential to build algorithms that learn with clinical care providers. To realize the potential of XAI, we must understand how each type of algorithm might fit into the real-world process of care delivery and the minds of medical decision-makers. At least initially, this will challenge algorithm developers to

understand clinical information and clinicians to efficiently integrate algorithms into their workflow.

Acknowledgments

The authors acknowledge the support of the NorthShore University Health System in providing the time, encouragement, and article processing fee (APF) funding to complete this manuscript. The authors also thank Clara Baumgarten in the Biomedical Visualization department of the University of Illinois at Chicago for producing a custom table of contents image in support of displaying the manuscript on the JMIR website.

Data Availability

All data generated or analyzed during this study are included in this published article.

Authors' Contributions

TG contributed clinical expertise and major ideas for the manuscript, wrote and influenced several sections of the paper, helped edit the paper, and compiled materials and visualizations. J Kang contributed clinical expertise for the manuscript, leadership for the project, and perspectives into applications of artificial intelligence in pathology. TT contributed to visualizations in the manuscript, added ideas for application of technology in clinical pathology, and edited the manuscript. J Krive contributed major ideas, literature review, provided clinical informatics and artificial intelligence expertise, compiled materials and supporting visualizations, helped edit the paper, and oversaw development of the manuscript.

Conflicts of Interest

J Kang is employed by Abbott Laboratories in their Tranafusion Medicine business unit. TG is employed by Fenwal, a Fresenius Kabi company. The knowledge shared in the manuscript is not influenced by any of these companies. All the other authors declare no conflicts of interest.

References

1. Farley B, Clark W. Simulation of self-organizing systems by digital computer. Trans IRE Prof Gr Inf Theory 1954 Sep;4(4):76-84. [doi: 10.1109/TIT.1954.1057468]
2. Friedman DJ, Parrish RG, Ross DA. Electronic health records and US public health: current realities and future promise. Am J Public Health 2013 Sep;103(9):1560-1567. [doi: 10.2105/AJPH.2013.301220] [Medline: 23865646]
3. Islam M, Hasan M, Wang X, Germack H, Noor-E-Alam M. A systematic review on healthcare analytics: application and theoretical perspective of data mining. Healthcare (Basel) 2018 May 23;6(2):54 [FREE Full text] [doi: 10.3390/healthcare6020054] [Medline: 29882866]
4. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Dec;2(4):230-243 [FREE Full text] [doi: 10.1136/svn-2017-000101] [Medline: 29507784]
5. Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: Current applications and future directions. PLoS Med 2018 Nov;15(11):e1002707 [FREE Full text] [doi: 10.1371/journal.pmed.1002707] [Medline: 30500815]
6. Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW, et al. Artificial intelligence in pathology. J Pathol Transl Med 2019 Jan;53(1):1-12 [FREE Full text] [doi: 10.4132/jptm.2018.12.16] [Medline: 30599506]
7. Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. Eur Radiol Exp 2018 Oct 24;2(1):35 [FREE Full text] [doi: 10.1186/s41747-018-0061-6] [Medline: 30353365]
8. Goebel R, Chander A, Holzinger K. Explainable AI: the new 42? In: CD-MAKE 2018: Machine Learning and Knowledge Extraction. 2018 Presented at: International Cross-Domain Conference for Machine Learning and Knowledge Extraction; August 27-30, 2018; Hamburg, Germany p. 295-303 URL: https://openaccess.city.ac.uk/id/eprint/20659/1 [doi: 10.1007/978-3-319-99740-7_21]
9. Yu K, Kohane I. Framing the challenges of artificial intelligence in medicine. BMJ Qual Saf 2019 Mar;28(3):238-241. [doi: 10.1136/bmjqs-2018-008551] [Medline: 30291179]
10. Holzinger A. From machine learning to explainable AI. In: DISA 2018 - IEEE World Symposium on Digital Intelligence for Systems and Machines, Proceedings. In: DISA 2018 - IEEE World Symposium on Digital Intelligence for Systems and Machines, Proceedings. IEEE; 2018 Presented at: 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA); August 23-25, 2018; Košice, Slovakia p. 55-66 URL: https://ieeexplore.ieee.org/abstract/document/8490530 [doi: 10.1109/disa.2018.8490530]
11. Tacchella A, Romano S, Ferraldeschi M, Salvetti M, Zaccaria A, Crisanti A, et al. Collaboration between a human group and artificial intelligence can improve prediction of multiple sclerosis course: a proof-of-principle study. F1000Res 2017;6:2172 [FREE Full text] [doi: 10.12688/f1000research.13114.2] [Medline: 29904574]

12. Miller D, Brown E. Artificial intelligence in medical practice: The question to the answer? Am J Med 2018 Feb;131(2):129-133 [FREE Full text] [doi: 10.1016/j.amjmed.2017.10.035] [Medline: 29126825]
13. Li D, Kulasegaram K, Hodges B. Why we needn't fear the machines: opportunities for medicine in a machine learning world. Acad Med 2019 May;94(5):623-625. [doi: 10.1097/ACM.0000000000002661] [Medline: 30768470]
14. 21 U.S.C. 360j - General provisions respecting control of devices intended for human use. U.S. Government Publishing Office. URL: https://tinyurl.com/555t6b4f [accessed 2023-08-18]
15. A guide to good practice for digital and data-driven health technologies. UK Department of Health and Social Care. URL: https://tinyurl.com/32srzb9t [accessed 2023-08-18]
16. Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges. N Engl J Med 2018 Mar 15;378(11):981-983 [FREE Full text] [doi: 10.1056/NEJMp1714229] [Medline: 29539284]
17. Gunning D. Explainable artificial intelligence. Defense Advanced Research Projects Agency. URL: https://www.darpa.mil/program/explainable-artificial-intelligence [accessed 2019-03-27]
18. Holzinger A, Biemann C, Pattichis C, Kell D. What do we need to build explainable AI systems for the medical domain? arXiv. Preprint posted online Dec 28, 2017. [FREE Full text] [doi: 10.1093/oso/9780197529003.003.0015]
19. Holzinger A, Malle B, Kieseberg P. Towards the pathologist: challenges of explainable-AI in digital pathology. arXiv. Preprint posted online on December 18, 2017. [FREE Full text]
20. Aminololama-Shakeri S, López JE. The doctor-patient relationship with artificial intelligence. AJR Am J Roentgenol 2019 Feb;212(2):308-310. [doi: 10.2214/AJR.18.20509] [Medline: 30540210]
21. Park S, Han K. Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. Radiology 2018 Mar;286(3):800-809. [doi: 10.1148/radiol.2017171920] [Medline: 29309734]
22. van der Hulle T, Cheung W, Kooij S, Beenen LFM, van Bemmel T, van Es J, YEARS study group. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. Lancet 2017 Jul 15;390(10091):289-297. [doi: 10.1016/S0140-6736(17)30885-1] [Medline: 28549662]
23. Qaseem A, Alguire P, Dallas P, Feinberg LE, Fitzgerald FT, Horwitch C, et al. Appropriate use of screening and diagnostic tests to foster high-value, cost-conscious care. Ann Intern Med 2012 Jan 17;156(2):147-149 [FREE Full text] [doi: 10.7326/0003-4819-156-2-201201170-00011] [Medline: 22250146]
24. Weng S, Reps J, Kai J, Garibaldi J, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? PLoS One 2017;12(4):e0174944 [FREE Full text] [doi: 10.1371/journal.pone.0174944] [Medline: 28376093]
25. Komorowski M, Celi L, Badawi O, Gordon A, Faisal A. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med 2018 Nov;24(11):1716-1720 [FREE Full text] [doi: 10.1038/s41591-018-0213-5] [Medline: 30349085]
26. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. arXiv. Preprint posted online on February 28, 2017. [FREE Full text]

Abbreviations

AI: artificial intelligence
DARPA: US Defense Advanced Research Projects Agency
XAI: explainable artificial intelligence

Edited by E Mensah; submitted 30.03.20; this is a non–peer-reviewed article;accepted 27.02.23; published 01.09.23.

Please cite as:
Gniadek T, Kang J, Theparee T, Krive J
Framework for Classifying Explainable Artificial Intelligence (XAI) Algorithms in Clinical Medicine
Online J Public Health Inform 2023;15:e50934
URL: https://ojphi.jmir.org/2023/1/e50934
doi:10.2196/50934
PMID:

Review

The Health Impact of mHealth Interventions in India: Systematic Review and Meta-Analysis

Vibha Joshi[1], PhD; Nitin Kumar Joshi[2,3], PhD; Pankaj Bhardwaj[2,3], MD; Kuldeep Singh[1], MD; Deepika Ojha[1], BTECH; Yogesh Kumar Jain[2,3], MPH

[1]Resource Centre Health Technology Assessment, All India Institute of Medical Sciences Jodhpur, Jodhpur, India

[2]School of Public Health, All India Institute of Medical Sciences Jodhpur, Jodhpur, India

[3]Department of Community Medicine and Family Medicine, All India Institute of Medical Sciences Jodhpur, Jodhpur, India

Corresponding Author:
Vibha Joshi, PhD
Resource Centre Health Technology Assessment
All India Institute of Medical Sciences Jodhpur
B110 Krishna Nagar
Basni-I
Jodhpur, 342005
India
Phone: 91 8290163030
Email: drvibhajoshi@gmail.com

Abstract

Background: Considerable use of mobile health (mHealth) interventions has been seen, and these interventions have beneficial effects on health and health service delivery processes, especially in resource-limited settings. Various functionalities of mobile phones offer a range of opportunities for mHealth interventions.

Objective: This review aims to assess the health impact of mHealth interventions in India.

Methods: This systematic review and meta-analysis was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Studies conducted in India, and published between April 1, 2011, and March 31, 2021, were considered. A literature search was conducted using a combination of MeSH (Medical Subject Headings) terms in different databases to identify peer-reviewed publications. Thirteen out of 1350 articles were included for the final review. Risk of bias was assessed using the Risk of Bias 2 tool for RCTs and Risk Of Bias In Non-randomised Studies - of Interventions tool (for nonrandomized trials), and a meta-analysis was performed using RevMan for 3 comparable studies on maternal, neonatal, and child health.

Results: The meta-analysis showed improved usage of maternal and child health services including iron–folic acid supplementation (odds ratio [OR] 14.30, 95% CI 6.65-30.75), administration of both doses of the tetanus toxoid (OR 2.47, 95% CI 0.22-27.37), and attending 4 or more antenatal check-ups (OR 1.82, 95% CI 0.65-5.09). Meta-analysis for studies concerning economic evaluation and chronic diseases could not be performed due to heterogeneity. However, a positive economic impact was observed from a societal perspective (ReMiND [reducing maternal and newborn deaths] and ImTeCHO [Innovative Mobile Technology for Community Health Operation] interventions), and chronic disease interventions showed a positive impact on clinical outcomes, patient and provider satisfaction, app usage, and improvement in health behaviors.

Conclusions: This review provides a comprehensive overview of mHealth technology in all health sectors in India, analyzing both health and health care usage indicators for interventions focused on maternal and child health and chronic diseases.

Trial Registration: PROSPERO 2021 CRD42021235315; https://tinyurl.com/yh4tp2j7

(Online J Public Health Inform 2023;15:e50927) doi:10.2196/50927

KEYWORDS

mobile applications; mobile apps; cost-benefit analysis; telemedicine; technology; India; patient satisfaction; pregnancy

Introduction

The use of mobile computing and communication technologies in health care and public health are seen as a rapidly expanding area within eHealth. The World Health Organization's Global Observatory for eHealth defined mobile health (mHealth) as "medical and public health practice supported by mobile devices, like mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices" [1]. Devices used in mHealth interventions include laptops, tablets, mobile phones, smartphones, palmtops, notebooks, and netbooks.

Features of mobile technology, including mobility, instantaneous access, and direct communication, permit faster transfer of health information, which aid in medical and public health practices. mHealth services range from simple apps to complex technologies including voice messaging, SMS text messaging, multimedia message service, Bluetooth technology, and others, which could transform the worldwide delivery of health services, especially in low- and middle-income countries [1].

Various functionalities such as SMS text messaging, voice messaging, mobile internet browsing, Voice over Internet Protocol services (eg, Skype), instant messaging services, photographic capabilities, and a wide variety of device-based applications available through mobile technology offer a range of opportunities for mHealth interventions, such as text message and interactive voice response campaigns and content to mobile phone–based imaging (which have potential diagnostic capabilities) [2,3]. This technology has a broad extent and accessibility, which can be efficiently leveraged for health care delivery in areas where access is a major constraint [4].

mHealth is increasingly being used for medical services and public health practice for patient communication, monitoring, and education [5,6]. The interventions have also shown to reduce the burden of diseases linked with poverty and an improvement in the accessibility of the health services in terms of clinical diagnosis, treatment adherence, and chronic disease management [1,7-9]. There is considerable interest in mHealth interventions with an enormous potential for beneficial effects on health and health service delivery processes, especially in resource-limited settings such as India [10].

This paper provides a review of evidence regarding the health impacts of mHealth interventions in India. The purpose of this review is to assess health impact in terms of measurable changes in mortality, morbidity, disability-adjusted life years (DALYs), and improved disease detection rates.

Methods

Study Design

This systematic review and meta-analysis was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analyses) guidelines [11]. Randomized controlled trials (RCTs), non-RCTs (including cluster RCTs and quasi-experimental studies), and prospective parallel cohort studies conducted in India were included. Studies published between April 1, 2011, and March 31, 2021, were considered, and the search was initiated on September 10, 2020, until March 10, 2021. Studies reported in the English language and conducted in India, which addressed the impact of mobile technology, using SMS text messaging or cellular telephone interventions for any disease (eg, diabetes, hypertension, cardiovascular disease, chronic respiratory disease, and cancer) and maternal and child health, and measured outcomes including morbidity, mortality, hospitalization rates, behavioral or lifestyle changes, the process of care improvements, clinical outcomes, patient and provider satisfaction, compliance, and cost-effectiveness, were included in the review.

Literature Search

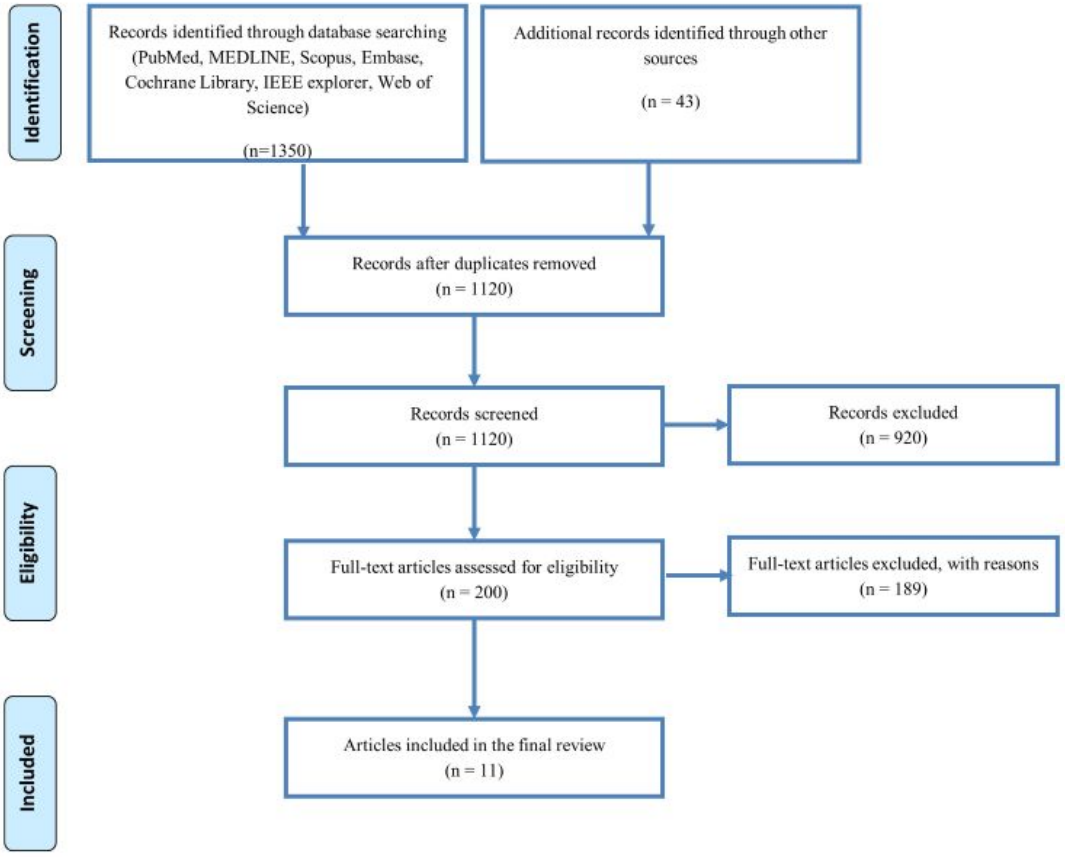
A literature search was conducted using a combination of text and Medical Subject Headings (MeSH) keywords in major databases, including PubMed, MEDLINE, Scopus, Cochrane Library, Web of Science, and Google scholar, to identify peer-reviewed publications. The MeSH keywords included the following: *Text Messaging*, *Health Literacy*, *Mobile Applications*, *Smartphone*, *Cell phone*, *Health Impact Assessment*, *Developing Countries*, *Multimedia*, *Cell Phone*, *Telemedicine*, *Medication Adherence*, *India*, *Hypertension*, *Primary Health Care*, *Risk Reduction Behavior*, *healthcare cost*, *Health Information Management*, and *Information Systems*. The search field was limited to the title or abstract (or both), and the type of publication was limited to original articles or full-length research articles. We excluded cross-sectional studies, letters, case reports, study protocols, reviews, opinions, gray literature, and non–peer-reviewed publications. The reference lists of articles were also examined to identify other potentially relevant articles. The protocol for this systematic review and meta-analysis has been registered in PROSPERO 2021 (CRD42021235315).

Study Selection and Characteristics

Two researchers (VJ and DO) independently screened the titles and abstracts to identify potentially eligible studies, and further assessment was performed by 2 authors (NKJ and YKJ). Only full-text articles published between 2011 and 2021, written in the English language, were included. The authors excluded duplicates and studies conducted outside India.

Initial searches identified 1393 titles. After removing duplicates, 1120 articles were included for initial screening. Of these, 920 articles were excluded after screening by title and abstract, leaving 200 articles, which were considered in more detail. A further 187 papers were subsequently excluded for not meeting the relevant criteria. Thirteen of the eligible studies were intervention studies, comprising 3 RCTs; 5 quasi–RCTs; 1 cluster RCT; 1 prospective, parallel-group cohort study; and 1 quantitative, single-arm, pretest, posttest interventional study (Figure 1).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for database searches of studies on mobile health interventions conducted in India in 2011-2020.



Data Extraction

The extracted data included the names of the authors, year of publication, study design, study location, sampling, and main results. All these details were captured and recorded in an Excel (Microsoft Corp) spreadsheet. The information reported in or calculated from the included studies was used for analysis. Corresponding authors of the articles were not contacted for unpublished or additional information. Disagreements related to the inclusion of an article were resolved through consensus among the authors.

Quality Assessment and Assessment of Risk of Bias

Risk of bias of each study was assessed using the Risk of Bias 2 tool for RCTs and Risk Of Bias In Non-randomised Studies - of Interventions for non-RCTs [12,13]. Risk-of-bias grading for the different components of each study is shown in Table 1. Four of the intervention studies were graded as being at low risk of bias, 6 as moderate, and 1 as high.

Table 1. Characteristics and results of studies investigating the effectiveness of mobile health (mHealth) interventions in India during 2011-2020.

Study (year; location) [overall risk of bias]	Study population	Intervention and control	Results	Study outcome
Prinja et al [14] (2017; Uttar Pradesh, India) [low]	• Population: data obtained from the 2011 AHS[a] and 2015 CEAHH[b] survey among women or mothers with 1-year-old children • Preintervention: 1508 ASHAs[c] (intervention: n=99; control: n=99); postintervention: 1028 (intervention: n=534; control: n=534)	• Intervention: pregnant women and mothers using an mHealth app; control: women and mothers not using mHealth applications	• Increase in the coverage IFA[d] supplementation (12.58%; 95% CI 0.086-0.27) • Self-reporting of illnesses or complication during pregnancy (13.11%) and after delivery (19.6%) • The coverage of ≥3 ANC[e] visits (10.3%; 95% CI 0.039-0.98) • Coverage of ≥2 tetanus toxoids (4.28%; 95% CI 0.055-0.68) • Institutional delivery (95% CI 0.044-0.59) • Full immunization (95% CI 0.20-1.032) • No change in the quality of ANC care	• Significant improvement in IFA supplementation, identification, and self-reporting of illnesses during pregnancy and after delivery
Modi et al [15] (Gujarat, India) [low]	• Population: rural tribal communities of Gujarat, India (neonates and mothers); population: 22 PHC[f] clusters (intervention: n=11; control: n=11)	• Intervention (with an mHealth package): 11 PHCs and 280 ASHAs; population: 234,134 • Control (without an mHealth package): 11 PHCs and 281 ASHAs; population: 242,809	• ANC of ≥4: intervention (n=622, 79.2%); 89.5, 95% 87.6-91.3); control (88.7, 95% CI 86.6-90.6) • TT[g] during the last pregnancy: intervention (n=771, 98.2%; 98.2, 95% CI 97.4-98.9); control (n=694, 98.3%; 96.8, 95% CI 96-97.6) • Delivered at an institution or hospital: intervention (n=580, 73.9%; 83.2, 95% CI 80.4-85.9); control (n=600, 85.0%; 84.9, 95% CI 82.1-87.6) • ASHAs present during delivery: intervention (n=267, 34.0%); control (n=267, 37.8%) • MACCI[h]: intervention (31%); control (31%) • ASHA visit at home at least twice in the first week of delivery: intervention (n=149, 19.0%; 32.4, 95% CI 29.7-35.1); control (n=99, 14.0%; 22.9, 95% CI 20.2-25.6) • Low Birth Weight (≤2 kg) at the time of birth: intervention (3.5, 95% CI 2.3-4.7); control (6.6; 95% CI 5.4-7.8) • Practice breastfeeding at 6 months: intervention (n=151, 19.2%; 57.4, 95% CI 54.1-60.8); control (n=95, 13.5%; 45.1, 95% CI 41.8-48.4)	• ImTeCHO[i] mobile apps and web-based applications, ASHAs, and PHC staff improved the coverage and quality of MNCH[j] services in difficult-to-reach areas • Improvement in coverage home visits by ASHAs during the antenatal period, postnatal period, early initiation of breastfeeding, and exclusive breastfeeding

Study (year; location) [overall risk of bias]	Study population	Intervention and control	Results	Study outcome
Murthy et al [16] (Mumbai, India) [moderate]	• 2016 pregnant women, aged 18 years or older: intervention (n=500); control (n=1516); analyzed (intervention: n=1038; control: n=379); time 1 (intervention: n=1516; control: n=500); time 2 (intervention: n=1113; control: n=402); time 3 (intervention: n=1038; control: n=379)	• Intervention group received mMitra voice messages twice per week throughout their pregnancy and until their infant turned 1 year of age • Control group received no mMitra voice message	• Infant care practices that the intervention group performed better: infant feeding at 6 months of age (OR[k] 1.4, 95% CI 1.08-1.82; *P*=.009), fully immunizing the infant (OR 1.531, 95% CI 1.141-2.055; *P*=.005) • Control group performed better on practices: increase in baby weight within 3 months (*P*=.03; OR 0.77, 95% CI 0.6-0.98) • In infant care knowledge: increase in baby solid food by 6 months in the intervention group (OR 1.89, 95% CI 1.371-2.605; *P*<.01); the ideal birth weight is >2.5 kg (OR 2.279, 95% CI 1.617-3.213; *P*<.01)	• mMitra voice-based mHealth intervention to demonstrate a positive impact on infant birth weight—a health outcome of public health importance
Ilozumba et al [17] (2018; Jharkhand, India) [low]	• Population: women between the ages of 18 and 45 years who had delivered a baby in the past 1 year (N=2200; intervention: n=733; control: n=739)	The study has 3 groups, all of which received standard care government programs that included the recruitment and support of ASHAs: • An intervention group that received MfM[l] in addition to an NGO's[m] existing interventions • A quasi-control group that received NGO programs • A standard care group that only received standard care government programs	• The odds of having a higher score on maternal health knowledge significantly increased when comparing intervention and control groups • Women in the MfM group were more likely to attend 4 or more ANC visits than those in the standard care group (OR 1.36, 95% CI 1.30-1.42) and the NGO group (OR 1.23, 95% CI 1.17-1.29) • The odds of a women in the MfM group were significantly higher than the odds of women in the standard care group (OR 1.34, 95% CI 1.28-1.41) and the NGO group (OR 1.19, 95% CI 1.13-1.25) • Higher maternal health knowledge -MfM versus standard care (intervention: OR 1.19, 95% CI 1.13-1.25; control [reference] OR 1.00) • Attended 4 or more ANC visits (intervention: OR 1.38, 95% CI 1.32-1.44; control [reference] OR 1.00) • Delivered at a health facility (intervention OR 1.35, 95% CI 1.29-1.42)	• This study showed that women in the intervention group reported higher levels of maternal health knowledge than those in the NGO intervention or those who received standard care • The primary outcomes of interest were maternal health knowledge, ANC attendance, and delivery in a health facility
Prinja et al [18] (2018; Uttar Pradesh, India) [low]		• Intervention: pregnant women and mothers using an mHealth app; control: women and mothers not using mHealth applications	• ReMiND[n] resulted in a cost saving of US $90 per DALY[o] averted US $2569 per death averted. From the health system perspective, ReMiND incurred an incremental cost of 12,993 (US $205) per DALY averted and 371,577 (US $5865) per death averted	• mHealth intervention as part of the ReMiND program is cost-saving from a societal perspective

Study (year; location) [overall risk of bias]	Study population	Intervention and control	Results	Study outcome
	• Population: data obtained from the 2011 AHS and 2015 CEAHH survey among women or mothers with 1-year-old children • Preintervention: 1508 ASHAs (intervention: n=99; control: n=99); postintervention: 1028 (intervention: n=534; control: n=534)			
Modi et al [19] (2020; Gujarat, India) [low]	• Population: rural tribal communities of Gujarat, India (neonates and mothers; population: N=22 PHC clusters: intervention: n=11; control: n=11)	• Intervention (with an mHealth package): 11 PHCs and 280 ASHAs; population: n=234,134 • Control (without an mHealth package): 11 PHCs and 281 ASHAs; population: n=242,809)	• ImTeCHO is a cost-effective intervention at an incremental cost of US $74 per life years saved or US $5057 per death averted • Total births in the study area (n=3014) • Cost per live birth (US $54) • Cost per 1000 live births (US $54,360) • Infant deaths averted per 1000 live births (n=11) • Life years saved (life expectancy=68.35 years; n=735) • Cost per infant deaths averted (US $5057) • Cost per life years saved due to infant deaths averted (US $74) • IMR^p as intention-to-treat in the study area (cost per ASHA (US $578.95)	• mHealth intervention as part of the ImTeCHO program is cost-effective and should be considered for replication
Pfammatter et al [20] (2015; India) [moderate]	• Population: adults aged 18 years and older (N=1925; intervention: n=611; control: n=632)	• Intervention: 1 million Nokia subscribers who opted into mDiabetes for 6 months • Control: non-Nokia phone subscribers	• Intervention group: 24.71% of them improved their fruit and vegetable intake and reduced their fat intake; 128 (20.95%) improved their preventive behavior • Control group: 36.55% decline in the number of participants' healthy behaviors; 73 (11.55%) improved their preventive behavior	• A text messaging intervention was feasible and showed initial evidence of effectiveness in improving diabetes-related health behaviors
Kleinman et al [21] (2017; India) [low]	• Population: aged 18-65 years with type 2 diabetes 6 months from baseline (N=90; intervention: n=44; control: n=46)	• Intervention: participants received the mHealth app and a mobile phone data stipend for 6 months • Control: manage their diabetes as usual		• Significantly more participants in the intervention group than in the control group

Study (year; location) [overall risk of bias]	Study population	Intervention and control	Results	Study outcome
			• Primary outcome: intervention mean 1.1-1.5; control mean 0.8-1.6 (*P*=.02) • Secondary outcomes: intervention mean 32.6-66.4; control mean 23.5-70.0 (*P*=.55) • BMI change: intervention mean 0.1-1.0; control mean 0.1-1.1 (*P*=.53)—patient-reported values improved from baseline to 6 months (intervention: n=16, 39.0%; control: n=5, 12.8%; *P*=.03) • Medication adherence (intervention: 39.0%; control: 12.8%; *P*=.03) • Increased frequency of blood glucose self-testing (intervention: 39.0%; control: 10.3%; *P*=.01)	
Prabhakaran et al [22] (2019; India) [moderate]	• Population: rural population (CHCs[q]), ≥30 years of age, confirmed diagnosis of hypertension or diabetes mellitus • Population: 40 clusters; intervention: n=20 (mWellcare) 20 clusters; 1842 participants enrolled (N=2140); control: 20 CHCs (allocated to EUC[r]) and 20 clusters; 1856 participants enrolled (N=2130)	• Effectiveness of the mWellcare app for 5 chronic conditions (hypertension, diabetes mellitus, current tobacco and alcohol use, and depression) vs usual care (intervention group: [mWellcare arm]: EUC NCD[s] nurses with the mWellcare system; control group: EUC NCD nurses Without the mWellcare system)		• Incremental benefit of mWellcare over enhanced usual care in chronic conditions • The trial did not find any significant difference in the primary outcomes, that is, reduction in SBP or HbA1c, and Secondary outcomes, that is, fasting blood glucose, total cholesterol, predicted 10-year risk of CVD, BMI, depression, and tobacco and alcohol use between the 2 arms

Study (year; location) [overall risk of bias]	Study population	Intervention and control	Results	Study outcome

			• Primary outcomes:
			  • Change in SBP[t]: control: mean –12.7 mm Hg; intervention: mean –13.7 mm Hg (effect size –0.3, adjusted 95% CI –3.9 to 3.3; *P*=.87)
			  • Change in HbA1c[u]: control: mean –0.58%; intervention: –0.48% (effect size 0.08, adjusted 95% CI –0.27 to 0.44; *P*=.66)
			• Secondary outcomes:
			  • Change in fasting blood glucose: control: mean –22.7 mg/dL; intervention: –15.0 mg/dL (effect size 8.4, adjusted 95% CI –9.6 to 26.5; *P*=.37)
			  • Change in total cholesterol: control: mean 2.0 mg/dL; intervention: mean 0.1 mg/dL (effect size –2.5, adjusted 95% CI –7.1 to 2.0; *P*=.29)
			  • Change in CVD[v] risk score: control: mean 0.6%; intervention: 2.4% (effect size –0.4, adjusted 95% CI –2.3 to 1.5; *P*=.66)
			  • Change in BMI: control: mean 0.08 kg/m2; intervention: 0.16 kg/m2 (effect size –0.05, adjusted 95% CI –0.47 to 0.37; *P*=.82)
			  • Change in tobacco use: control: mean –7.0%; intervention: mean –0.6% (effect size –0.8, adjusted 95% CI –5.7 to 4.2; *P*=.76)
			  • Change in alcohol use: control: mean –3.8%; intervention: mean –2.4% (effect size 0.7, adjusted 95% CI –3.7 to 5.1; *P*=.74)
			  • Change in alcohol use score: control: mean 10.0; intervention: 9.4 (effect size –0.6, adjusted 95% CI –3.2 to 2.1; *P*=.68)
			  • Change in depression score: control: mean 12.4; intervention: mean 10.9 (effect size –1.6, adjusted 95% CI –4.4 to 1.2; *P*=.28)	

Study (year; location) [overall risk of bias]	Study population	Intervention and control	Results	Study outcome
Garner et al [23] (2020; India) [moderate]	• Population: urban slum and rural slum dwellers (n=346) • Pretest (n=87): those who earned an 8 or above on the pretest paired t test • Posttest (n=259): those who earned a 7 or below on the pretest	• Intervention through an mHealth app to improve hypertension health literacy	• Study aim 1: to assess the effectiveness of an mHealth app to improve hypertension health literacy among participants in India • Study aim 2: to estimate relationships between participant hypertension health literacy and sociodemographic variables • Pretest: participants who performed moderately well on the pretest also had improved posttest scores (significant mean difference between pretest and posttest scores 2.49; *P*<.001 [paired t test])	• The mHealth app provides an effective and valuable culturally tailored educational resource for nurses and other health to improve hypertension health literacy among populations in India
Gautham et al [24] (2015; Tamil Nadu, India) [high]	• Population: rural health providers (n=16) and patients (n=126; experimental: n=65; control: n=61)	• Intervention group: given applications on their mobile phones • Control group: no application given; only the phone and a set of paper guidelines to use in the field	• Control group scored significantly higher than the experimental group (control group: mean 13.68; experimental group: 9.51; *P*<.05) in the posttraining evaluation. • Control: mean pretraining score 8.58 out of 19 (SD 2.03); experimental: mean pretraining score 7.01 out of 19 (SD 1.85; *P*=.19) • Control: mean posttraining score 13.68 out of 19 (SD 2.17); experimental: mean posttraining score 9.51 out of 19 (SD 2.48; *P*<.05)	• This study supports the implication that mMRIGs[w] comprise a feasible and effective solution for standardizing and enhancing the quality of care delivered by millions of frontline rural health providers with varying levels of training and literacy
Praveen et al [25] (2014; Andhra Pradesh, India) [moderate]	• Population: ASHAs, NPHWs[x], and PHC physicians. 227 adults screened by ASHAs, 65 adults screened by PHC physicians	• The CDSS[y] was field-tested in 11 villages and 3 PHCs. CVD risk factor profile for participants screened by ASHAs (n=227) and doctors (n=65)	• The CDSS recommend referral to a doctor to 128 of 227 adults and did not recommend referral to 99 of 227 adults. • High CVD risk was noted in 88 of 128 (69%) adults, and in another 40 of 99 (31%) adults. • Blood pressure lowering medication given to 29 of 65 (45%) adults and not to 36 of 65 (55%) adults. • The other assessment of behavior change (COM-B[z] model) revealed 3 themes: (1) potential to transform prevailing health care models, (2) task-shifting of CVD screening to the ASHA was the central driver of change, and (3) system-level barriers such as access to doctors and medicines are still present	• A tablet-based CDSS implemented within primary health care systems has the potential to help improve CVD outcomes in India
Jadhav et al [26] (2016; Maharashtra, India) [moderate]				• Reinforcement of oral health education through SMS text messages is effective media to improve oral health

Study (year; location) [overall risk of bias]	Study population	Intervention and control	Results	Study outcome
	• Population: adults aged 18-20 years having a personal mobile phone with SMS text messaging capability (N=400; control: n=200; intervention: n=200)	• Intervention group: the message was reinforced through SMS text messages from mobile phones • Control: no oral health–related SMS text messages or any kind of health education was given to the participants	• Gender-wise distribution of participants: 137 male and 63 female participants in the intervention group and 149 male and 51 female participants in the control group (*P*>.05) • Mean OHI[aa] score at different intervals between the intervention and control groups showed no significant difference at baseline (*P*=.28) and after the first month (*P*=.58); however, it was significantly lower in the intervention group after the second, third, and sixth months (*P*<.01) • Mean GI[ab] scores at different intervals between the intervention and control groups were significantly no different at baseline (*P*=.39) and after the first month (*P*=.85); however, it was significantly lower in the intervention group after the second, third, and sixth months (*P*<.01)	

[a]AHS: Annual Health Survey.

[b]CEAHH: cost-effectiveness analysis household.

[c]ASHA: accredited social health activist.

[d]IFA: iron–folic acid.

[e]ANC: antenatal care.

[f]PHC: primary health center.

[g]TT: Tetanus toxoid.

[h]MACCI: modified accredited social health activist–centric composite coverage index.

[i]ImTeCHO: Innovative Mobile Technology for Community Health Operation.

[j]MNCH: maternal, neonatal, and child health.

[k]OR: odds ratio.

[l]MfM: Mobile for Mothers.

[m]NGO: nongovernmental organization.

[n]ReMiND: reducing maternal and newborn deaths.

[o]DALY: disability-adjusted life year.

[p]IMR: infant mortality rate.

[q]CHC: community health center.

[r]EUC: enhanced usual care.

[s]NCD: noncommunicable disease.

[t]SBP: systolic blood pressure.

[u]HbA_{1c}: hemoglobin A_{1c}.

[v]CVD: cardiovascular disease.

[w]mMRIG: media-rich interactive guideline.

[x]NPHW: nonphysician health care worker.

[y]CDSS: clinical decision support system.

[z]COM-B: capability, opportunity, and motivation.

[aa]OHI: Oral Hygiene Index.

[ab]GI: Gingival Index.

Meta-Analysis

There was substantial heterogeneity among studies in their mHealth interventions and outcomes, except for studies on maternal, neonatal, and child health. Consequently, we performed a random-effects meta-analysis using the Mantel-Haenszel method in RevMan [27] for 3 comparable studies, which had all used cell phones rather than routine prenatal care as the intervention and had assessed increases in the number of antenatal check-ups, tetanus toxoids administered to pregnant women, institutional deliveries, and iron–folic acid to assess the effect of health care usage. However, as the relevant intervention for the purpose of this review, we exclusively compared the cell phone group to the usual care group in the meta-analysis. However, given the small number of studies, we did not undertake possible sensitivity analyses.

Results

Types of Outcomes Examined

Four studies examined the indicators of maternal, neonatal, and child health [14-17]—these reported the number of antenatal check-ups [14,15,17]; birth weight [15]; institutional delivery [14-17]; knowledge of the danger signs of pregnancy [14,15]; indicators of infant feeding and breastfeeding [14]; usage of antenatal, intrapartum, and postnatal care [14,15,17]; indicators of self-efficacy [15,17]; uptake of immunization [14,15]; and maternal health knowledge [17]. We found two studies evaluating the cost-effectiveness of mHealth programs [18,19]. Other outcomes included improvement in diabetes risk behaviors and increased awareness about the causes and complications of diabetes [20], improvement in medication adherence and the frequency of blood glucose testing [21], change in systolic blood pressure and hemoglobin A_{1c} levels [22], quality of care delivered by primary health workers [23-25], and oral health education [26]. The results are organized below in accordance with the types of outcomes examined in each study.

Effects on Maternal, Neonatal, and Child Health

A pre-post quasi-experimental study used an mHealth application in the Kaushambi district in Uttar Pradesh, India, to increase the quality of counseling by community health volunteers, resulting in improved uptake of maternal, neonatal, and child health services. A significant increase in coverage iron–folic acid supplementation and identification and self-reporting of illnesses or complications during pregnancy and after delivery were seen in the intervention area, but there was no change in the quality of antenatal care (ANC) care [14]. Similarly, an mHealth application was used in an open cluster RCT conducted in 22 primary health centers in 6 tribal blocks of Bharuch and Narmada districts in Gujrat, India, to assess the increase in the coverage of maternal, neonatal, and child health services and that of at least 2 home visits by accredited social health activists within the first week of birth. There were significant improvements in coverage home visits by accredited social health activists during the antenatal and postnatal period, early initiation of breastfeeding, and exclusive breastfeeding [15].

A pseudo-RCT conducted in Mumbai (Maharashtra, India) by Murthy et al [16], assessed the impact of age- and stage-based mobile phone voice messaging for pregnant women on reduction in low birth weight and child malnutrition and improvement in women's infant care knowledge and practices. They observed that the intervention group performed well in infant care practice indicators: administering supplementary feeding to the infant at 6 months of age (odds ratio [OR] 1.4, 95% CI 1.08-1.82; *P*=.009) and fully immunizing the infant (OR 1.531, 95% CI 1.141-2.055; *P*=.005). Moreover, women in the intervention group had increased knowledge of giving infants solid food by 6 months of age and of the fact that the ideal birth weight is >2.5 kg [16]. A study from Jharkhand used a mobile app to support home visits by community health workers; Ilozumba et al [17] found that women receiving the mHealth intervention were more likely to attend 4 or more ANC visits and had significantly higher odds of delivering a baby at a health center than those receiving standard care and those receiving other interventions from a nongovernmental organization. Moreover, the usage of ANC services and delivery at a health center were associated with the education level of the spouse [17].

Cost-Effectiveness

Prinja et al [18] assessed the cost-effectiveness of the ReMiND (reducing maternal and newborn deaths) program in Uttar Pradesh, India; both the societal and health care perspectives were taken into account. Overall, the ReMiND program was considered a cost-saving intervention from the societal perspective. It resulted in a cost saving of US $90 per DALY averted US $2569 per death averted. From the health system perspective, the ReMiND program incurred an incremental cost of 12,993 (US $205) per DALY averted and 371,577 (US $5865) per death averted [18]. A study conducted in Gujrat, India, found the ImTeCHO (Innovative Mobile Technology for Community Health Operation) intervention to be cost-effective at an incremental cost of US $74 per life-years saved or US $5057 per death averted [19].

Effect on Chronic Conditions

Study conducted by Pfammatter et al [20] to examine the effect of mDiabetes—a text messaging program to improve diabetes risk behaviors—on fruit, vegetable, and fat intake and exercise among Nokia phone users in India. A greater improvement in the health behavior composite score over 6 months was observed among participants who received the text messages than among those who did not receive text messages [20]. An RCT conducted by Kleinman et al [21] at 3 sites in India assessed the impact of an mHealth diabetes platform on clinical outcomes, patient-reported outcomes, patient and provider satisfaction, and app usage. There was decrease of 1.5% in mean hemoglobin A_{1c} levels in the intervention group and 0.8% in the usual care group, an improvement in self-reported medication adherence from baseline, and an increase in blood glucose testing in the intervention group from baseline compared to that in the control group (39.0% vs 10.3%, respectively; *P*=.01) [21]. Prabhakaran et al [22] conducted a cluster-RCT using the mWellcare system for integrated management of 5 chronic conditions (hypertension, diabetes mellitus, current tobacco and alcohol use, and depression). No evidence of

difference in systolic blood pressure and hemoglobin A_{1c} levels was observed between the intervention and control groups [22].

Other Effects

Garner et al [23] determined the effectiveness of an mHealth application to improve hypertension health literacy among vulnerable populations in India. A significant improvement in the understanding of hypertension through the innovative animated application was observed [23]. In the RCT conducted in rural areas of Tamil Nadu, India, Gautham et al [24] observed that mobile app–based procedural guidance for rural frontline health care providers had significant potential for attaining consistently standardized quality of care with patients' acceptance. Praveen et al [25] showed that implementation of a mobile clinical decision support system for cardiovascular disease management by public nonphysician health care workers and physicians in a rural Indian setting increased the number of referrals to the physician and had potential to help improve cardiovascular disease outcomes, but system-level barriers have an impact on limiting the access to medical care. Jadhav et al [26] assessed the effectiveness of the reinforcement of oral health education SMS text messages and reported that mean Oral Hygiene Index and Gingival Index scores in the intervention group were significantly lower than those in the control group (P<.01).

Effect on Health Care Usage

Among pregnant women, those using mHealth interventions were more likely to take a complete dose of iron–folic acid supplements (OR 14.30, 95% CI 6.65-30.75; Figure 2), both doses of the tetanus toxoid (OR 2.47, 95% CI 0.22-27.37; Figure 3), and to attended 4 or more antenatal care check-ups (OR 1.82, 95% CI 0.65-5.09; Figure 4) than those who received routine prenatal care. No strong evidence of differences regarding institutional deliveries (OR 1.14, 95% CI 0.26-4.95) were found.

Figure 2. Meta-analysis of the effect of mobile health interventions versus standard care on the intake of complete doses of iron–folic acid supplements during prenatal care. IFA: iron–folic acid; mHealth: mobile health; OR: odds ratio; SC: standard care.

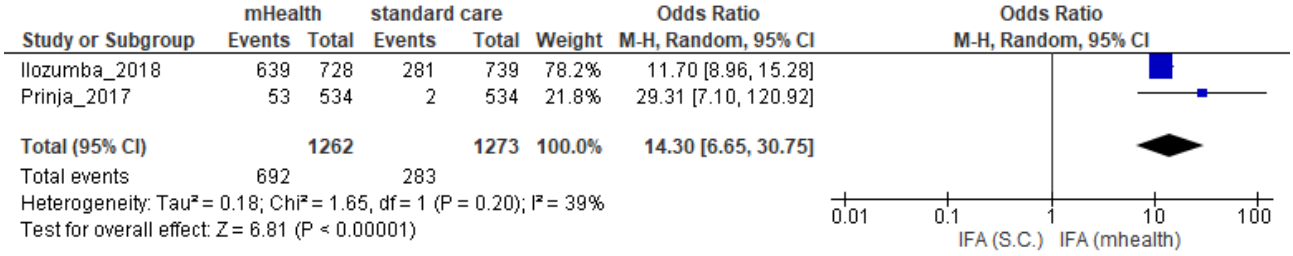


Figure 3. Meta-analysis of the effect of mobile health interventions versus standard care on taking 2 doses of the tetanus toxoid during pregnancy. mHealth: mobile health; OR: odds ratio; TT: tetanus toxoid.

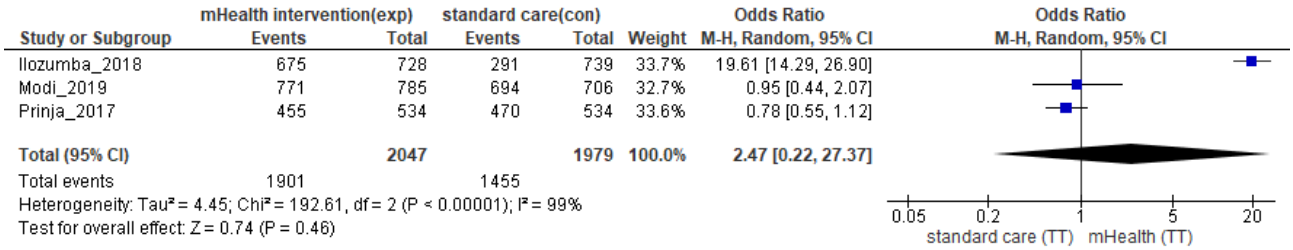
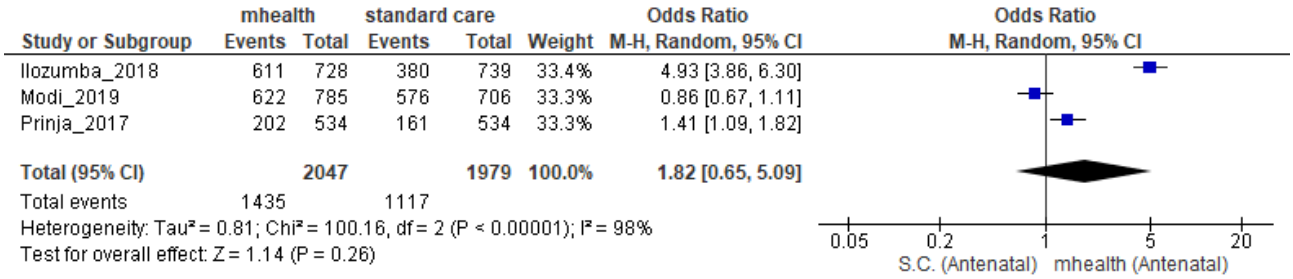


Figure 4. Meta-analysis of the effect of mobile health interventions versus standard care on 3 or more antenatal care check-ups conducted during pregnancy. mHealth: mobile health; OR: odds ratio; SC: standard care.



Discussion

Principal Findings

mHealth is an implicit, promising tool for addressing several health care system limitations in transitional countries, such as a limited health care workforce, scarce resources, high burden of disease, rapid population growth, and challenges of extending health care to underserved populations. We identified 13 studies showing the impact of mobile technology–based interventions designed to improve health care service delivery processes in the Indian setting. Most studies were at moderate and low risk of bias. Heterogeneity among studies did not allow the calculation of a pooled estimate for all the parameters. However,

a meta-analyses of 3 studies arbitrated to be sufficiently homogenous showed that mHealth interventions used for maternal and child health improved the usage of prenatal services including the intake of a complete dose of iron–folic acid supplements, taking both doses of the tetanus toxoid, and attending 4 or more antenatal care check-ups. No strong evidence of differences regarding institutional deliveries were found. A similar review conducted by Lee et al [28] for low- to middle-income countries showed that mHealth technologies are rapidly being used to promote health care use, improve the quality of pre- and postnatal care, and collect data on pregnancy and child health.

In our systematic review, we could not use economic evaluation–tailored reporting standards (such as the CHEERS [Consolidated Health Economic Evaluation Reporting Standards] checklist [29]) for full economic evaluation due to the lack of sufficient economic evaluation studies, as indicated by Iribarren et al [30], who described the evidence related to economic evaluations of mHealth interventions in low- to middle-income countries and in the evaluation of 2 mHealth interventions in India: ReMiND [18] and ImTeCHO [19]. These studies included a comparison of the effectiveness of a health-related outcome and reported economic data. Both the studies showed a positive economic impact considering the societal perspective.

All the studies included in this review provide evidence that the interventions conducted for the chronic diseases had an impact on clinical outcomes, patient and provider satisfaction, app usage, and improvement in health behavior (except for the study conducted by Prabhakaran et al [22]). Similar findings were described in the review conducted by Beratarrechea et al [31] for chronic diseases in transitional countries, which addressed more than 1 outcome and reported a positive impact on chronic disease outcomes.

Limitations and Conclusion

This paper reviews the comprehensive use of mHealth technologies in all sectors of health care in India. We used a thorough, extensive, and highly sensitive literature search technique in this systematic review, which analyses both health and health care usage indicators, encompassing the entire scope of relevant mHealth technologies including those focusing on maternal and child health and chronic diseases. All comparative reviews have been conducted for low- to middle-income countries and mainly focused on the either chronic disease or maternal and child health [28,30-38].

However, due to a small number of studies for a single set of interventions, a meta-analysis for all the impact indicators was not conducted. Additional work is needed to improve and test this with a larger set of interventions, and to determine how to best integrate it with different conceptual frameworks that have been published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File , 32 KB - ojphi_v15i1e50927_app1.docx]

References

1. World Health Organization. MHealth: New Horizons for Health Through Mobile Technologies. Geneva: World Health Organization; 2011.
2. Devi BR, Syed-Abdul S, Kumar A, Iqbal U, Nguyen P, Li YJ, et al. mHealth: An updated systematic review with a focus on HIV/AIDS and tuberculosis long term management using mobile phones. Comput Methods Programs Biomed 2015 Nov;122(2):257-265. [doi: 10.1016/j.cmpb.2015.08.003] [Medline: 26304621]
3. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug 06;1(2):160-171. [doi: 10.9745/ghsp-d-13-00031]
4. Goel S, Bhatnagar N, Sharma D, Singh A. Bridging the human resource gap in primary health care delivery systems of developing countries with mHealth: narrative literature review. JMIR Mhealth Uhealth 2013 Dec 03;1(2):e25 [FREE Full text] [doi: 10.2196/mhealth.2688] [Medline: 25099436]
5. mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World. United Nations Office for the Coordination of Humanitarian Affairs. 2009. URL: https://reliefweb.int/attachments/7816247a-20a6-3322-a7ee-2e6f824ec8cf/A695C439809709CF49257689001019B9-UNF_Feb2009.pdf [accessed 2023-08-21]
6. Carter A, Liddle J, Hall W, Chenery H. Mobile phones in research and treatment: ethical guidelines and future directions. JMIR Mhealth Uhealth 2015 Oct 16;3(4):e95 [FREE Full text] [doi: 10.2196/mhealth.4538] [Medline: 26474545]
7. Carter A, Liddle J, Hall W, Chenery H. Mobile phones in research and treatment: ethical guidelines and future directions. JMIR Mhealth Uhealth 2015 Oct 16;3(4):e95 [FREE Full text] [doi: 10.2196/mhealth.4538] [Medline: 26474545]
8. Donner J. Research approaches to mobile use in the developing world: a review of the literature. Inf Soc 2008 May 06;24(3):140-159. [doi: 10.1080/01972240802019970]

9. GSMA Development Fund, The Cherie Blair Foundation for Women, Vital Wave Consulting. Women and Mobile: A Global Opportunity. University of Birmingham. 2010. URL: https://gsdrc.org/document-library/women-and-mobile-a-global-opportunity/ [accessed 2021-06-06]
10. Nglazi MD, Bekker L, Wood R, Hussey GD, Wiysonge CS. Mobile phone text messaging for promoting adherence to anti-tuberculosis treatment: a systematic review. BMC Infect Dis 2013 Dec 02;13(1):566 [FREE Full text] [doi: 10.1186/1471-2334-13-566] [Medline: 24295439]
11. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
12. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898 [FREE Full text] [doi: 10.1136/bmj.l4898] [Medline: 31462531]
13. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919 [FREE Full text] [doi: 10.1136/bmj.i4919] [Medline: 27733354]
14. Prinja S, Nimesh R, Gupta A, Bahuguna P, Gupta M, Thakur JS. Impact of m-health application used by community health volunteers on improving utilisation of maternal, new-born and child health care services in a rural area of Uttar Pradesh, India. Trop Med Int Health 2017 Jul 13;22(7):895-907 [FREE Full text] [doi: 10.1111/tmi.12895] [Medline: 28510997]
15. Modi D, Dholakia N, Gopalan R, Venkatraman S, Dave K, Shah S, et al. mHealth intervention "ImTeCHO" to improve delivery of maternal, neonatal, and child care services-A cluster-randomized trial in tribal areas of Gujarat, India. PLoS Med 2019 Oct;16(10):e1002939 [FREE Full text] [doi: 10.1371/journal.pmed.1002939] [Medline: 31647821]
16. Murthy N, Chandrasekharan S, Prakash MP, Kaonga NN, Peter J, Ganju A, et al. The impact of an mHealth voice message service (mMitra) on infant care knowledge, and practices among low-income women in India: findings from a pseudo-randomized controlled trial. Matern Child Health J 2019 Dec 04;23(12):1658-1669 [FREE Full text] [doi: 10.1007/s10995-019-02805-5] [Medline: 31584144]
17. Ilozumba O, Van Belle S, Dieleman M, Liem L, Choudhury M, Broerse JEW. The effect of a community health worker utilized mobile health application on maternal health knowledge and behavior: a quasi-experimental study. Front Public Health 2018;6:133 [FREE Full text] [doi: 10.3389/fpubh.2018.00133] [Medline: 29868541]
18. Prinja S, Bahuguna P, Gupta A, Nimesh R, Gupta M, Thakur JS. Cost effectiveness of mHealth intervention by community health workers for reducing maternal and newborn mortality in rural Uttar Pradesh, India. Cost Eff Resour Alloc 2018 Jun 25;16(1):25 [FREE Full text] [doi: 10.1186/s12962-018-0110-2] [Medline: 29983645]
19. Modi D, Saha S, Vaghela P, Dave K, Anand A, Desai S, et al. Costing and cost-effectiveness of a mobile health intervention (ImTeCHO) in improving infant mortality in tribal areas of Gujarat, India: cluster randomized controlled trial. JMIR Mhealth Uhealth 2020 Oct 14;8(10):e17066 [FREE Full text] [doi: 10.2196/17066] [Medline: 33052122]
20. Pfammatter A, Spring B, Saligram N, Davé R, Gowda A, Blais L, et al. mHealth intervention to improve diabetes risk behaviors in India: a prospective, parallel group cohort study. J Med Internet Res 2016 Aug 05;18(8):e207 [FREE Full text] [doi: 10.2196/jmir.5712] [Medline: 27496271]
21. Kleinman NJ, Shah A, Shah S, Phatak S, Viswanathan V. Improved medication adherence and frequency of blood glucose self-testing using an m-Health platform versus usual care in a multisite randomized clinical trial among people with type 2 diabetes in India. Telemed J E Health 2017 Sep;23(9):733-740. [doi: 10.1089/tmj.2016.0265] [Medline: 28328396]
22. Prabhakaran D, Jha D, Prieto-Merino D, Roy A, Singh K, Ajay VS, Members of the Research Steering Committee,Investigators,Members of the Data SafetyMonitoring Board. Effectiveness of an mHealth-based electronic decision support system for integrated management of chronic conditions in primary care: the mWellcare cluster-randomized controlled trial. Circulation 2019 Jan 15;139(3):380-391 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.118.038192] [Medline: 30586732]
23. Garner S, George C, Young P, Hitchcock J, Koch H, Green G, et al. Effectiveness of an mHealth application to improve hypertension health literacy in India. Int Nurs Rev 2020 Dec 07;67(4):476-483. [doi: 10.1111/inr.12616] [Medline: 32767464]
24. Gautham M, Iyengar MS, Johnson CW. Mobile phone-based clinical guidance for rural health providers in India. Health Informatics J 2015 Dec 12;21(4):253-266 [FREE Full text] [doi: 10.1177/1460458214523153] [Medline: 24621929]
25. Praveen D, Patel A, Raghu A, Clifford GD, Maulik PK, Mohammad Abdul A, et al. SMARTHealth India: development and field evaluation of a mobile clinical decision support system for cardiovascular diseases in rural India. JMIR Mhealth Uhealth 2014 Dec 08;2(4):e54 [FREE Full text] [doi: 10.2196/mhealth.3568] [Medline: 25487047]
26. Jadhav H, Dodamani A, Karibasappa G, Naik R, Khairnar M, Deshmukh M, et al. Effect of reinforcement of oral health education message through short messaging service in mobile phones: a quasi-experimental trial. Int J Telemed Appl 2016;2016:7293516 [FREE Full text] [doi: 10.1155/2016/7293516] [Medline: 26941793]
27. Review M(CP. Version 5. 3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration. 2014. URL: https://training.cochrane.org/handbook/archive/v6.2/chapter-10 [accessed 2021-06-22]
28. Lee SH, Nurmatov UB, Nwaru BI, Mukherjee M, Grant L, Pagliari C. Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: Systematic review and meta-analysis. J Glob Health 2016 Jun;6(1):010401 [FREE Full text] [doi: 10.7189/jogh.06.010401] [Medline: 26649177]

29. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, ISPOR Health Economic Evaluation Publication Guidelines-CHEERS Good Reporting Practices Task Force. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health 2013 Mar;16(2):231-250 [FREE Full text] [doi: 10.1016/j.jval.2013.02.002] [Medline: 23538175]
30. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? A systematic review of economic evaluations of mHealth solutions. PLoS One 2017;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
31. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. Telemed J E Health 2014 Jan;20(1):75-82 [FREE Full text] [doi: 10.1089/tmj.2012.0328] [Medline: 24205809]
32. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? Glob Health Action 2014 Oct 27;7(1):25606 [FREE Full text] [doi: 10.3402/gha.v7.25606] [Medline: 25361730]
33. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. J Health Commun 2012 May 02;17 Suppl 1(sup1):82-104. [doi: 10.1080/10810730.2011.649160] [Medline: 22548603]
34. Peiris D, Praveen D, Johnson C, Mogulluru K. Use of mHealth systems and tools for non-communicable diseases in low- and middle-income countries: a systematic review. J Cardiovasc Transl Res 2014 Nov 11;7(8):677-691. [doi: 10.1007/s12265-014-9581-5] [Medline: 25209729]
35. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
36. Beatty AL, Fukuoka Y, Whooley MA. Using mobile technology for cardiac rehabilitation: a review and framework for development and evaluation. JAHA 2013 Nov 18;2(6). [doi: 10.1161/jaha.113.000568]
37. Bloomfield GS, Vedanthan R, Vasudevan L, Kithei A, Were M, Velazquez EJ. Mobile health for non-communicable diseases in Sub-Saharan Africa: a systematic review of the literature and strategic framework for research. Global Health 2014 Jun 13;10(1):49 [FREE Full text] [doi: 10.1186/1744-8603-10-49] [Medline: 24927745]
38. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]

Abbreviations

ANC: antenatal care
CHEERS: Consolidated Health Economic Evaluation Reporting Standards
DALY: disability-adjusted life year
ImTeCHO: Innovative Mobile Technology for Community Health Operation
MeSH: Medical Subject Headings
mHealth: mobile health
OR: odds ratio
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
ReMiND: reducing maternal and newborn deaths

Edited by E Mensah; submitted 26.02.22; peer-reviewed by VJ Latha Bommu, P Boakye; accepted 03.02.23; published 04.09.23.

Please cite as:
Joshi V, Joshi NK, Bhardwaj P, Singh K, Ojha D, Jain YK
The Health Impact of mHealth Interventions in India: Systematic Review and Meta-Analysis
Online J Public Health Inform 2023;15:e50927
URL: https://ojphi.jmir.org/2023/1/e50927
doi:10.2196/50927
PMID:

Editorial

Completion of the Transfer of the Online Journal of Public Health Informatics (OJPHI) to JMIR Publications

Edward Mensah[1], MPhil, PhD

School of Public Health, University of Illinois Chicago, Chicago, IL, United States

Corresponding Author:
Edward Mensah, MPhil, PhD
School of Public Health
University of Illinois Chicago
1603 West Taylor Street
Chicago, IL, 60612
United States
Phone: 1 7735926958
Email: dehasnem@uic.edu

Abstract

Founded in 2009, the *Online Journal of Public Health Informatics* (OJPHI) strives to provide an unparalleled experience as the platform of choice to advance public and population health informatics. As a premier peer-reviewed journal in this field, OJPHI's mission is to serve as an advocate for the discipline through the dissemination of public health informatics research results and best practices among practitioners, researchers, policymakers, and educators. However, in the current environment, running an independent open access journal has not been without challenges. Judging from the low geographic spread of our current stakeholders, the overreliance on a small volunteer management staff, the limited scope of topics published by the journal, and the long article turnaround time, it is obvious that OJPHI requires a change in direction in order to fully achieve its mission. Fortunately, our new publisher JMIR Publications is the leading brand in this field, with a portfolio of top peer-reviewed journals covering innovation, technology, digital medicine and health services research in the internet age. Under the leadership of JMIR Publications, OJPHI plans to expand its scope to include new topics such as precision public health informatics, the use of artificial intelligence and machine learning in public health research and practice, and infodemiology in public health informatics.

(Online J Public Health Inform 2023;15:e50243) doi:10.2196/50243

KEYWORDS

public health informatics; data science; precision public health; artificial intelligence; health promotion; disease prevention

Introduction

The *Online Journal of Public Health Informatics* (OJPHI) has been delivering the latest developments in the emerging field of public health informatics since 2009. The journal was originally created to fill a gap in the public health informatics publishing and training landscape. I started to recognize the need for such a journal as a result of my own involvement in advancing training and education in public health informatics. In 2002, I had the privilege to cofound and serve as director of the graduate program in Public Health Informatics at the University of Illinois Chicago (UIC) School of Public Health, the very first program of its kind in the United States.

The immeasurable value of transforming raw data into information and knowledge for effective and efficient decision-making by using information and communication technologies (ICTs) was evident during the COVID-19 pandemic. Although public health is an information-intensive field, it lags behind other health-related fields in the utilization of ICTs for the delivery of services and resource management. The emergence of public health informatics as a professional specialty is part of a larger development of informatics in health-related fields, including medicine, nursing, pharmacy, and dentistry. The interest in informatics as a specialty within these fields reflects the significance of data collection, analysis, and transformation into information and knowledge in the health care sector. Several journals have been launched in response to the growing need for informaticians in these health care disciplines, and many of the JMIR Publications journals, not least one of their flagship journals *JMIR Public Health and Surveillance*, also publish research at the intersection of public health and technology.

Unfortunately, the COVID-19 pandemic and the subsequent increased interest and research output in our discipline also demonstrated the limitations of a self-published, open access journal relying largely on volunteer efforts. Not only is it

becoming increasingly challenging to find peer reviewers, but it is also increasingly difficult to operate in a rapidly changing and complex scholarly publications landscape while running on a shoestring budget. With a broader shift toward open access, libraries now often direct their publications budget toward transformative (hopefully transitional!) agreements that support costs to publish in former subscription journals of large publishing corporations and phasing out institutional open access funds, leaving little or no support for independent journals.

Considering all these factors, it is now time to put the future of OJPHI into professional hands, by choosing a mission-driven publisher that has its roots in academia. JMIR Publications, a medium-sized but rapidly growing publisher with its mission-driven academic leadership and focus on innovation in health and medicine, is the ideal fit as a new home and publisher of OJPHI. With access to the experienced professional staff of JMIR Publications, OJPHI will attract stakeholders from a wide variety of disciplines and geographic areas, achieve a shorter turnaround time, and upgrade the quality of the journal. The new publisher also has the capacity to make deals with institutional partners to put the journal on a more financially stable basis.

Like other journals in the JMIR Publications portfolio, OJPHI seeks to promote interdisciplinary collaboration and welcomes contributions by researchers and practitioners from a wide range of fields, including public health, computer science, data science, health informatics, and related disciplines. As such, it is complementary to the over 34 other titles in the JMIR Publications portfolio, and authors will have the opportunity to transfer their submissions between journals without the need for a new peer review (portable peer reviews).

OJPHI invites submissions of original research articles, reviews, and perspectives or viewpoints that cover a broad range of topics related to public health informatics, including the following:

- Use of health ICTs and data science to improve public health
- Development and implementation of electronic health records and other health information systems by public health agencies
- Framework, evaluation, and use of health information exchange technologies among public health agencies, hospitals, laboratories, and clinics
- Use of data analytics (including artificial intelligence [AI], machine learning, geographic information system, visualization, and data mining technologies) in public health research and practice
- Use of social media in public and population health informatics applications for health promotion and disease prevention
- Evaluation of mobile health technology and digital platforms in public health practice
- Integration of social determinants of health within public health practice using ICTs
- Ethical, legal, and social implications of public health informatics
- Precision public health informatics
- Analysis and utilization of big data for health promotion and health equity
- Development and evaluation of contact tracing technologies for public health practice and policy
- Infodemiology in public health informatics (complementing similar sections in the *Journal of Medical Internet Research*, *JMIR Public Health and Surveillance*, and the recently launched journal *JMIR Infodemiology*)

The manuscript management system and journal homepage has been migrated to the JMIR Publications platform (under its new URL ojphi.jmir.org) and submissions are now open as of July 2023. Articles under consideration on the current platform are in the process of being migrated, and the review process for the articles in the pipeline will commence thereafter. For new articles, authors are encouraged to indicate their preferences by selecting from the list the topics under which they want their articles to be reviewed. This will aid search engine optimization and make it easier to automate the process of allocating the articles to specific reviewers. Authors of manuscripts submitted to other JMIR Publications journals are also encouraged to consider OJPHI as a destination for manuscript transfer.

In addition, we are refreshing our editorial board, and I invite interested researchers who are passionate about advancing the field of public health informatics to contact us. I am personally excited about the future of the field and that of the journal, which I consider in excellent hands with JMIR Publications. I thank our stakeholders, including authors and peer reviewers, for their continued patience and support as we are making this transition.

Joining JMIR Publications, an established and top-notch publisher, positions OJPHI within a wider ecosystem of open access biomedical and health informatics journals. We now have the facilities needed to reach out to a diverse group of authors and readers from universities, public health agencies, and policy makers in developed countries and the Global South. JMIR Publications provides us with the platform to improve the quality of the journal, article turnaround time, and become the public and population health informatics journal of choice among our stakeholders.

Conflicts of Interest

EM is the editor-in-chief of the *Online Journal of Public Health Informatics* (OJPHI) and receives a nominal honorarium.

Abbreviations

AI: artificial intelligence
ICT: information and communication technology
OJPHI: Online Journal of Public Health Informatics
UIC: University of Illinois Chicago

Edited by A Mavragani; submitted 23.06.23; this is a non–peer-reviewed article;accepted 26.06.23; published 18.07.23.

Please cite as:
Mensah E
Completion of the Transfer of the Online Journal of Public Health Informatics (OJPHI) to JMIR Publications
Online J Public Health Inform 2023;15:e50243
URL: https://ojphi.jmir.org/2023/1/e50243
doi:10.2196/50243
PMID:

Original Paper

Toxicology Test Results for Public Health Surveillance of the Opioid Epidemic: Retrospective Analysis

Titus Schleyer[1,2], DMD, PhD; Bill Robinson[3], BS; Samir Parmar[4], MCS, MPH; Diane Janowiak[3], BS; P Joseph Gibson[5], MPH, PhD; Val Spangler[3], MBA

[1]Center for Biomedical Informatics, Regenstrief Institute, Inc, Indianapolis, IN, United States

[2]School of Medicine, Indiana University, Indianapolis, IN, United States

[3]hc1 Insights, Inc, Indianapolis, IN, United States

[4]Atlanta, GA, United States

[5]Marion County Public Health Department, Indianapolis, GA, United States

Corresponding Author:
Titus Schleyer, DMD, PhD
Center for Biomedical Informatics
Regenstrief Institute, Inc
1101 W 10th St
Indianapolis, IN, 46202
United States
Phone: 1 3172749204
Email: schleyer@regenstrief.org

Abstract

Background: Addressing the opioid epidemic requires timely insights into population-level factors, such as trends in prevalence of legal and illegal substances, overdoses, and deaths.

Objective: This study aimed to examine whether toxicology test results of living individuals from a variety of sources could be useful in surveilling the opioid epidemic.

Methods: A retrospective analysis standardized, merged, and linked toxicology results from 24 laboratories in Marion County, Indiana, United States, from September 1, 2018, to August 31, 2019. The data set consisted of 33,787 Marion County residents and their 746,681 results. We related the data to general Marion County demographics and compared alerts generated by toxicology results to opioid overdose–related emergency department visits. Nineteen domain experts helped prototype analytical visualizations. Main outcome measures included test positivity in the county and by ZIP code; selected demographics of individuals with toxicology results; and correlation of toxicology results with opioid overdose–related emergency department visits.

Results: Four percent of Marion County residents had at least 1 toxicology result. Test positivity rates ranged from 3% to 19% across ZIP codes. Males were underrepresented in the data set. Age distribution resembled that of Marion County. Alerts for opioid toxicology results were not correlated with opioid overdose–related emergency department visits.

Conclusions: Analyzing toxicology results at scale was impeded by varying data formats, completeness, and representativeness; changes in data feeds; and patient matching difficulties. In this study, toxicology results did not predict spikes in opioid overdoses. Larger, more rigorous and well-controlled studies are needed to assess the utility of toxicology tests in predicting opioid overdose spikes.

(Online J Public Health Inform 2023;15:e50936) doi:10.2196/50936

KEYWORDS

opioid epidemic; clinical laboratory techniques; public health; epidemiology; toxicology

Introduction

A key challenge in addressing the opioid epidemic [1,2] is timely insight into population-level factors, such as trends in prevalence of legal and illegal substances, overdoses, and deaths. Many surveillance systems and approaches at the national, regional, and local levels exist [3-12] but are limited by being (1) focused on late-stage outcomes such as drug-related arrests and overdose deaths [13,14], (2) frequently not available until long after an

event occurs [15,16], (3) drawn from fragmented and siloed data, and (4) not representative [17].

This project explored whether toxicology laboratory results [18,19] from testing in health care and jail settings (in short, "toxicology results") are potentially useful in surveilling the opioid epidemic. Our work builds on similar efforts to leverage calls to poison control centers for surveillance [5]. These selected toxicology tests may be useful because they occur in a variety of settings where effects of changing opioid use may first become apparent, can be communicated in real time through health information technology standards such as Health Level 7 (HL7) [20] and Logical Observation Identifiers Names and Codes (LOINC) [21,22], can be integrated at an individual level using record linkage [23], and are required in many states' prescription drug monitoring programs and are recommended for monitoring patients on chronic opioid therapy [24,25].

To date, toxicology results have been used primarily for retrospective, one-off analyses [26-29]. The goal of this project was to investigate whether ongoing, timely monitoring of living individuals' toxicology results gathered from several sources might indicate changes in the general population's opioid use. This study addresses the following questions: (1) how can toxicology test data from multiple sources be aggregated and homogenized? (2) What are the characteristics of persons in an aggregated set of toxicology test data in Marion County, Indiana, United States, compared to those of the general population? (3) Can toxicology test data provide direct indicators for trends regarding the opioid epidemic? (4) How might such toxicology test data be integrated into a dashboard for managing the opioid epidemic?

Methods

Overview

The health care company managing this project receives many types of laboratory tests from its clients, which are primarily clinical laboratories. The orders for and results of these tests are transmitted to the health care company in near real time using the HL7 protocol. The transmissions typically contain patient demographics, ordering provider and location, specimen information, the ordered tests, and the quantitative or qualitative results. All laboratory test information is transmitted to a data lake where it is refined, enriched, and deidentified.

Sample Characteristics

The data set for this study included all toxicology results with a patient (or, if not available, an order or accession) address in Marion County (the largest county in Indiana with a population of 954,760 individuals as of 2018; home to Indianapolis—the capital of Indiana) collected between September 1, 2018, and August 31, 2019.

Patient and provider records were refined using several methods, including standardizing variable values and formatting, decomposing composite fields, and retrieving missing information (eg, ZIP codes based on address). We refined organizations through a similar process and categorized them by type, such as addiction treatment centers, criminal justice, forensics, hospital, emergency department, pain management, and primary care. Patient records were linked through a string similarity function that assigned a master patient identifier if records had a match rate of 95% or greater.

Matching tests across multiple laboratories was one of the most challenging aspects of cleaning and homogenizing the data. We used string matching functions and manual review to assign a LOINC code to each test, which we then mapped to a local drug class hierarchy.

For result records, abnormality was calculated by comparing the result value with the transmitted reference range. Positivity for toxicology tests, based on keywords or numeric values, were determined by profiling client data. Multiple tests for the same patient were considered as separate, with the exception of pairs of screening and confirmatory tests (which occurred rarely and were considered positive if the confirmatory test yielded a positive finding).

We only included records with ZIP codes from Marion County (either the patients' or, if unavailable, the ordering location's records). We retained only the data from hospitals, primary care providers, clinical specialty providers, and jails because emergency department, coroner, forensics, police department, sheriff's offices, state police, and employer testing data were expected to exhibit markedly different result patterns. For instance, in emergency departments and law enforcement, sampling due to suspected alcohol and drug use typically results in high positivity rates. Positivity rates for employment drug testing, on the other hand, are often low since individuals applying for jobs know a drug test is required. Because these patterns were observed in our data, we excluded results from these settings. We included data from jails because positivity rates were fairly consistent with those reported in Marion County.

In total, 24 clients of the health care company had data for at least 1 Marion County patient. The largest contributor provided 64% of the results' volume but only supplied data from January to April 2018. We excluded these data because they mostly comprised employment testing and had, comparatively, a much lower positivity rate. Of the remaining data, 67% of them were obtained from a regional reference laboratory and the core laboratory for several hospitals in Indiana, and the next 15% of them were obtained from a laboratory carrying out testing for law enforcement and forensics (only jail data were included). The remainder of the laboratories were primarily regional toxicology and reference laboratories. In addition to our data refinement and linking infrastructure, we already had built a preliminary dashboard for visualizing the data that served as the basis for this project [30].

Dashboard Development and Data Analysis

The project was advised by a 9-member external advisory group consisting of 3 academic researchers; 5 public health professionals at the local, state, and international level; and 1 corporate participant. This group met several times with 8 health care company staff members and executive leaders over the course of the project period to provide high-level strategic guidance. A technical working group, consisting of 3 members of the external advisory group and company technical personnel,

prepared and analyzed the data and designed and prototyped the dashboard.

After data preparation, we summarized toxicology results data descriptively and compared them to data for Marion County where possible. We performed 2-proportion *z* tests on each category, excluding unknown counts in totals. In addition, we developed a set of design ideas for a local dashboard to manage the opioid epidemic and evaluated them through a survey of the advisory group and additional company personnel. In the survey, we presented proposed design features and asked one or more questions, such as "What kind of useful information can you glean from the presented visualization?" "What kind of information is missing?" "Is it easy to determine values of interest?" The survey was distributed to 19 invitees (8 advisory group members and 11 company staff).

Last, we evaluated how toxicology results trends related to signals derived from opioid overdose–related emergency department visits. The goal of this analysis was to determine whether simple positivity rates from toxicology results can provide useful signals for trends regarding the opioid epidemic. For instance, intuition would suggest that test positivity rates might rise prior to spikes in overdoses. For the toxicology tests, we used the specimen collection date, and for emergency department encounters, the visit date.

The Marion County Department of Health uses ESSENCE (Electronic Surveillance System for the Early Notification of Community-Based Epidemics) [31] to analyze opioid-related data, such as opioid overdose–related emergency department visits [32], and generate alerts for notable events. To detect spikes in test positivity, or the incidence of opioid overdose–related emergency department visits, we applied the ESSENCE C2 detection method to toxicology results and emergency department opioid overdose data between September 1, 2018, and August 31, 2019. The algorithm uses a moving sample average and sample SD to standardize each observation, with a 2-day lag in the mean and SD calculations [33]. The implemented baseline in ESSENCE is 28 days, compared to the baseline of 7 days. For emergency department data, the opioid outbreak indicator was the daily count of individuals with any overdose, and for laboratory data, the daily positive proportion of opioid toxicology tests. If the result exceeded 3 SDs above the sample mean, an alert was generated.

Ethics Approval

This project (protocol #1802267756: Development and formative evaluation of the Opioid Epidemic Management Dashboard) was approved as expedited by the Indiana University institutional review board on February 2, 2018.

Results

Overview

Table 1 shows a comparison of the major characteristics of the health care company's data set and Marion County demographics. For the study period, 4% of people with a Marion County address had at least 1 test result. The health care company data set's gender distribution (35.9% males and 64.1% females—within the 82.5% of individuals with a known gender) differed significantly from the gender distribution of Marion County's population (48.2% males and 51.8% females). A much larger, national data set of test results had a more similar gender distribution (40.5% males and 59.5% females). In numerous records, data on race and ethnicity were missing and therefore not included. Age distributions (within the 81.9% of individuals with a known age) also showed differences, with individuals aged up to 19 years significantly underrepresented and those aged 20 to 39 years significantly overrepresented in the health care company data set.

Table 2 provides additional detail about toxicology results for the 37 ZIP codes in Marion County. The proportion of residents by ZIP code with at least 1 toxicology test result within the study period ranged from 0.4% to 41.5%. In 28 (76%) ZIP codes, the range was between 0.4% and 3%; in 5 (14%), between 5.1% and 8.4%; and in 4 (11%), between 10.7% and 41.5%. For the 46.2% of records having no patient address, the ZIP code of the ordering location was used, which implies that ZIP codes with large order volumes and low populations showed higher percentages of tested residents. For example, ZIP code 46202, which has the highest percentage at 41.5%, is the location of Marion County Jail II. The next 4 highest percentages are in ZIP codes that include major hospitals. The result positivity rate, defined as the number of positive results divided by the number of nonmissing or nondeterminate results, ranged from 3% to 19%. Visits to the emergency department due to overdose and overdose deaths are provided for context. However, it should be noted that the time periods for the number of residents and overdose deaths are for 2018, only partially overlapping with the September 2018 to August 2019 date range of the laboratory tests.

Table 1. Comparison of gender and age characteristics of the health care company's data set (September 1, 2018, to August 31, 2019) to Marion County demographics obtained from US Census Bureau (2018).

	Health care Company (N=33,787), n (%)	Marion County (N=954,670), n (%)	*P* value (*Z* test)
Gender			
Male	10,012 (35.9)	460,093 (48.2)	<.01
Female	17,856 (64.1)	494,577 (51.8)	<.01
Unknown	5923 (N/A[a])	N/A (N/A)	N/A
Age (years)			
0-19	3901 (14.1)	257,636 (27.0)	<.01
20-39	14,268 (51.5)	293,706 (30.8)	<.01
40-59	6046 (21.8)	228,542 (23.9)	<.01
60-79	3088 (11.2)	145,891 (15.3)	<.01
>80	391 (1.4)	28,895 (3)	<.01
unknown	6093 (N/A)	N/A (N/A)	N/A

[a]N/A: not applicable.

Table 2. Toxicology results for the 37 ZIP codes in Marion County, Indiana, United States, sorted by the number of residents in descending order from September 1, 2018, to August 31, 2019 (except for the number of residents and overdose deaths from January 1, 2018, to December 31, 2018).

ZIP code	Residents, n	Residents in the data set, n (%)	Residents with positive results, n	Results, n	Positive results, n	Result positivity rate (%)	Overdose-related emergency department visits, n	Overdose deaths, n
46227	56,449	2931 (5.2)	1545	43,211	4508	10.4	264	20
46226	45,998	1183 (2.6)	729	18,119	1842	10.2	208	14
46237	39,803	427 (1.1)	268	7312	934	12.8	128	13
46203	38,313	546 (1.4)	359	7958	1023	12.9	282	20
46254	36,530	934 (2.6)	424	17,121	1238	7.2	66	8
46224	35,177	575 (1.6)	331	9977	1382	13.9	93	<5
46220	33,833	751 (2.2)	439	11,509	1423	12.4	64	<5
46219	33,646	4407 (13.1)	3030	59,509	8048	13.5	269	17
46222	33,061	372 (1.1)	219	4774	534	11.2	153	19
46260	32,779	2768 (8.4)	1690	40,864	4746	11.6	79	<5
46241	30,918	420 (1.4)	288	3488	550	15.8	177	23
46218	30,516	903 (3.0)	587	12,892	1306	10.1	215	16
46201	30,487	863 (2.8)	639	13,216	1817	13.7	361	34
46217	29,558	296 (1.0)	165	4082	363	8.9	103	10
46235	29,507	766 (2.6)	387	11,117	817	7.3	99	5
46229	27,913	1437 (5.1)	1019	21,806	4182	19.2	105	13
46221	27,054	206 (0.8)	105	2988	295	9.9	146	7
46236	26,751	462 (1.7)	242	6522	546	8.4	74	<5
46268	26,411	602 (2.3)	342	8511	726	8.5	64	<5
46205	26,351	553 (2.1)	293	8037	668	8.3	86	11
46234	25,002	158 (0.6)	77	2299	180	7.8	34	7
46239	24,348	311 (1.3)	134	4594	304	6.6	91	5
46214	23,747	275 (1.2)	108	3984	243	6.1	46	<5
46256	23,541	3887 (16.5)	1847	54,212	4155	7.7	55	5
46208	23,312	290 (1.2)	181	4042	394	9.7	77	7
46240	18,817	374 (2.0)	230	4912	478	9.7	40	<5
46250	18,545	490 (2.6)	282	7271	683	9.4	28	5
46202	16,021	6652 (41.5)	5447	263,328	19,823	7.5	139	5
46113	15,037	72 (0.5)	27	1048	66	6.3	22	<5
46228	14,876	182 (1.2)	104	2570	220	8.6	36	<5
46107	12,801	110 (0.9)	49	1721	122	3.0	70	5
46259	11,777	61 (0.5)	25	782	62	7.9	30	<5
46231	11,440	47 (0.4)	24	681	68	10.0	19	<5
46278	7968	98 (1.2)	50	1509	102	6.8	7	<5
46225	6524	489 (7.5)	417	66,361	2014	3.0	65	8
46204	5903	369 (6.3)	291	37,083	1066	2.9	82	<5
46216	1496	160 (10.7)	97	2348	297	12.6	<5	<5
Total	932,210	35,427 (3.8)	22,491	771,758	67,225	8.7	N/A[a]	N/A

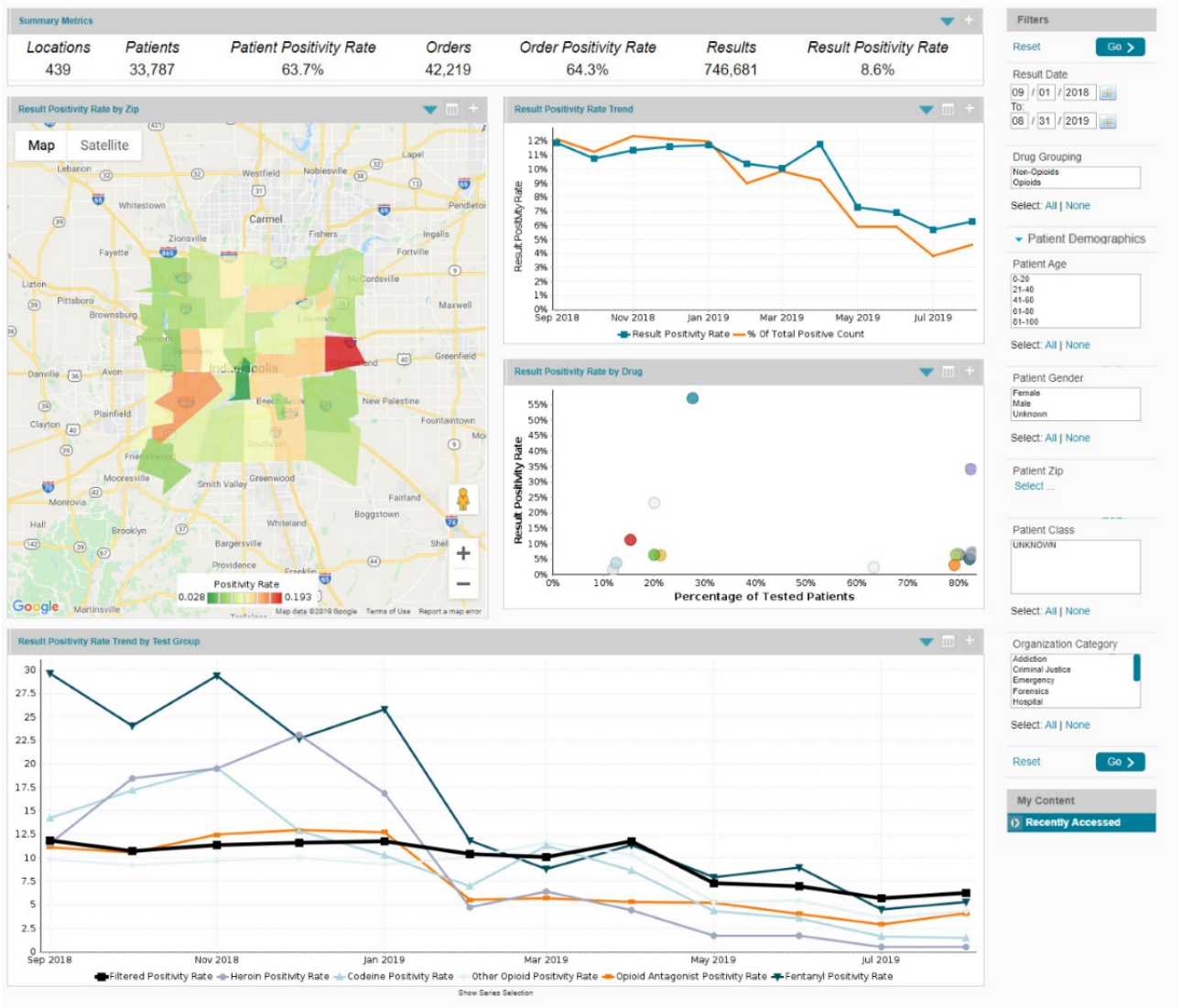
[a]N/A: not applicable.

Dashboard Design and Prototyping

We used the data to explore potential visualizations based on the dashboard we had previously developed [30]. We focused our efforts on designing specific enhancements to the dashboard and identifying potential improvements through a survey. Of the 19 invitees, 10 responded. A description of the main design enhancements and potential improvements follows.

Figure 1 shows the final design of the main dashboard. The Summary Metrics bar near the top summarizes the data in general with regard to patients, test results, and positivity. The result positivity rate by ZIP heat map provides a geographic overview of Marion County; other graphs display general and drug-specific positivity trends and information. More information can be displayed by hovering over certain areas of the screen. Filters on the right allow the user to subset the data. The following sections provide additional detail and survey results for selected aspects of the design.

Figure 1. Final design of main dashboard displaying summary toxicology results for Marion County, Indiana, United States.



Summary Metrics Display

The Summary Metrics display in Figure 1 shows the following key numbers:

- Locations: count of distinct physical ordering sites such as a doctor's office, emergency department, employer, or clinic
- Patients: total number of individual patients with 1 or more test results
- Patient positivity rate: number of patients with at least 1 positive result divided by the number of patients with 1 or more test results
- Orders: total number of unique laboratory orders including 1 or more results
- Order positivity rate: the number of orders with 1 or more positive result divided by the number of orders with 1 or more results
- Results: total number of results (screening and confirmatory tests for the same drug only counted once)
- Result positivity rate: number of positive results divided by the total number of results

Feedback on the Summary Metrics display included the need for clearer labeling of selected metrics, separating toxicology results for licit or illicit substances, and adding contextual data,

such as the total population from which the tests are drawn, naloxone administrations, fatal or nonfatal overdoses, overdose-related emergency medical services runs, prescriptions, and medication-assisted treatment volume.

Result Positivity Rate by ZIP Code Display

Geographical maps are common in displaying data related to the opioid epidemic [34]. The result positivity rate by ZIP display in Figure 1 shows the map displaying test positivity rates across ZIP codes. Positivity rates range from 0.028 to 0.193. The user can pan and zoom in or out of the map, as well as select data to display using the filters on the right of the dashboard.

Feedback on this design included its usefulness for identifying "hot spots," and the need to standardize the color range across displays with different minima and maxima of the positivity rate; providing the numerator and denominator for the positivity rate, as well as residents by ZIP code, to judge representativeness of the data; the ability to "scrub" through the time line; and the ability to correlate with other data, such as overdoses or emergency medical service runs. In addition, map areas did not correspond exactly with ZIP codes.

Drug Positivity by Age, Gender, and Drug

Figure 2 shows drug positivity by age, gender, and drug to understand multivariate relationships in the data. Certain patterns are evident, such as generally lower positivity rates for heroin in females than in males, and age and positivity differentials regarding cocaine.

Feedback on this design included that it was easy to tell which groups are high-risk and whether these groups were stable over time. It was perceived as difficult to tell how important or statistically significant the differences were between rectangles of different colors. An alternative design suggestion was a bar graph by age as an initial visual, with a drill-down option to look at time trends.

Figure 2. Visualizations for drug positivity by age, gender, and specific drug for Marion County, Indiana, United States. Data are from September 1, 2018, to August 31, 2019.

Toxicology Results as Predictive Signals

We used the ESSENCE platform to determine the potential relationship between changes in toxicology results' positivity and opioid-related chief complaints in emergency departments. We included toxicology results from our data set likely to be predictive for future overdoses, such as those generated in health care and jail facilities. We excluded data that were likely to have been collected after overdoses, such as emergency department and coroner data. During the 12 months of overlap between the data sets, ESSENCE generated 4 alerts for emergency department visits and 3 for toxicology results. We counted the combination of 1 alert each as an "episode" if (1) the toxicology result alert occurred prior to the emergency department visit alert and (2) both alerts occurred within a 30-day window but not on the same day. We chose 30 days as

the time window because we considered toxicology results alerts outside of that window as not actionable. Only 1 episode occurred in the data set, with the toxicology results alert preceding the emergency department alert by 9 days. The sample size was not sufficient to conduct statistical tests for comparison.

Discussion

The purpose of this project was to elucidate whether analyzing toxicology results may be useful in monitoring the opioid epidemic. Our key findings are summarized below.

Challenges in Aggregating, Preparing, and Managing Toxicology Test Data From Multiple Sources

We aggregated data from 24 health care company clients, all of whom sent us data in varying formats and degrees of completeness. Attempts to combine data sources for population surveillance need to account for differing formats, rates of completeness, and missingness.

Changes in the client base and data feeds affected data availability. Where possible, we imputed missing data by matching to a more complete version of the patient record from another client, or using addresses of ordering providers and accessioning location as proxies. However, such imputations carried the risk of introducing bias.

Linking data was relatively easy because data were fully identified and could be matched across clients. However, we could not link toxicology results to external data such as nonfatal overdoses, overdose deaths, or naloxone administrations due to privacy constraints. This limitation reduced our ability to develop a more complete picture of the epidemic.

Factors Increasing the Effort Required to Clean and Synthesize Data

Test names are often not standardized among laboratories, requiring significant computational or manual inferencing. While initial test mapping took considerable effort, we partially automated the process as the set of test names mapped to the hierarchy grew. Artificial intelligence methods using the training data generated in this project may facilitate test mapping in the future.

Toxicology test orders often include component tests for multiple drugs. The component tests, in turn, can have multiple instances such as screening (qualitative) and confirmatory (quantitative) tests for the same drug. Many confirmatory tests are a collection of metabolites that can indicate 1 or more parent drugs. We counted 1 or more positive results for the same drug within the same test order as a single result.

Toxicology results can sometimes be difficult to interpret with respect to the source substances introduced into the patient's system and the metabolites detected at various time points.

Representativeness of Toxicology Test Data for Larger Trends in the Opioid Epidemic

Toxicology tests are typically not administered to a random sample of the population. For instance, pain management patients are more likely to be tested when drug testing is required for chronic opioid therapy. Such consistency testing necessarily reflects the expectations of the clinician, such as a positive result when the patient is on opioid therapy. On the other hand, drug screening related to employment or Department of Transportation monitoring samples a different demographic with the expectation that most test results are negative. Inclusion or exclusion of data sets generated for various purposes will likely skew positivity rates. This may be partially addressed by only including data sources that are not likely to be strongly biased with regard to the test result, and weighting included data sources according to their demographic composition, to approximate the demographics of the population of interest. Encouragingly, the gender distribution in our results resembled that found in the results of a large, national laboratory test provider, providing some evidence of external validity. Unfortunately, demographic information is often missing in laboratory test records.

Potential Approaches to Visualizing Toxicology Test Data

Our project generated several potentially useful ideas for visualizing toxicology test data. Summary statistics that include unique individuals, the number of orders and tests, and positive or negative test results for various analytes could help monitor drug use or abuse prior to serious events, such as overdoses and overdose deaths. A variety of visualization techniques can help show relationships among and trends for selected variables.

However, limitations in being able to integrate and interlink different data sets was a key obstacle for generating insights. For instance, several of our organizational participants had access to highly relevant data, such as prescriptions, fatal and nonfatal overdoses, emergency medical service runs, and drug seizures related to opioids. Interlinking these data on an individual basis (where possible) was perceived as potentially useful but challenging with regard to governance, record linking, and time and effort required.

Using Toxicology Test Data as a "Signal" in Surveillance

Our results are inconclusive regarding the question of whether surveillance of toxicology results at the urban county level can serve as an effective predictor for spikes in opioid overdose cases admitted to emergency departments. While we focused on data likely to be predictive for such events, the proportion of overdose cases for which a toxicology test result is available prior to or after an overdose is unknown. In addition, chief complaints and discharge diagnoses for Marion County vary considerably by hospital with regard to specificity about overdoses and specific drugs involved. Individual-level data linkage may be a promising option to answer such questions and elicit more meaningful signals than possible in our study.

Scaling Our Approach to Other Municipalities and States for Public Health Surveillance Purposes

Due to variation in data set content, availability, granularity, and linkability, our approach is likely difficult to scale easily to other municipalities and states. The health care company's market position in Indiana provided a strong foundation for attempting to explore the utility of test results for tracking the

opioid epidemic. However, even given that, it is unknown to what degree toxicology results in Marion County are indicative of trends in the opioid epidemic.

Currently, toxicology tests among living individuals appear to play only a small role in surveillance of the opioid epidemic. However, such tests might be important and timely indicators for drug use disorder trends in the general population. Further work should address issues identified in our study, such as aggregating, preparing, and managing toxicology test data; representativeness of these data; potential approaches to visualizing them; and using toxicology test data as a "signal" in surveillance.

Conclusions

Analyzing toxicology results of living individuals from a variety of sources may be useful as an indicator of trends in opioid use. Important findings to consider include the following: (1) there are multiple challenges in aggregating, preparing, and managing toxicology test results for population trend analysis; (2) the representativeness of these data for the general population must be assessed carefully; (3) leveraging toxicology test results as a "signal" in surveillance likely requires robust data sets and sophisticated analyses. Individual-level data linkage may be a promising option to elicit more meaningful signals than is currently possible; and (4) a variety of visualization techniques can help show relationships among and trends for selected variables.

Acknowledgments

This work was supported by National Institutes of Health (NIH; award R43DA045455: "Combating the opioid epidemic with big data analytics and a live, nationwide dashboard") to hc1 Insights, Inc. DJ, BR, and VS are employees of hc1. During the performance period for the NIH award, VS was a part-time employee of hc1. Part of this publication was made possible by the Lilly Endowment, Inc. Physician Scientist Initiative and the Indiana Clinical and Translational Sciences Institute (funded in part by grant ULI TR002529) from the Clinical and Translational Science Award of the NIH's National Center for Advancing Translational Sciences. We gratefully acknowledge the contributions of Zach Berg, Brad Bostic, Laura Breedlove, Charlie Clark, Shandy Dearth, Marion Greene, Jim Huizenga, Scott LaNeve, Mark Preston, Brad Ray, Russ Ray, Heather Stith, and especially Shelly Simeone. SP's contributed to this work during his employment at the Marion County Public Health Department.

Conflicts of Interest

None declared.

References

1. Jalal H, Buchanich JM, Roberts MS, Balmert LC, Zhang K, Burke DS. Changing dynamics of the drug overdose epidemic in the United States from 1979 through 2016. Science 2018 Sep 21;361(6408) [FREE Full text] [doi: 10.1126/science.aau1184] [Medline: 30237320]
2. O'Donnell JK, Halpin J, Mattson CL, Goldberger BA, Gladden RM. Deaths involving fentanyl, fentanyl analogs, and U-47700 - 10 states, July-December 2016. MMWR Morb Mortal Wkly Rep 2017 Nov 03;66(43):1197-1202 [FREE Full text] [doi: 10.15585/mmwr.mm6643e1] [Medline: 29095804]
3. Addressing Prescription Drug Abuse in the United States: Current Activities and Future Opportunities. Department of Health and Human Services. URL: https://www.cdc.gov/drugoverdose/pdf/hhs_prescription_drug_abuse_report_09.2013.pdf [accessed 2023-09-07]
4. California Overdose Surveillance Dashboard. URL: https://skylab.cdph.ca.gov/ODdash/ [accessed 2020-03-18]
5. Davis JM, Severtson SG, Bucher-Bartelson B, Dart RC. Using poison center exposure calls to predict prescription opioid abuse and misuse-related emergency department visits. Pharmacoepidemiol Drug Saf 2014 Jan 16;23(1):18-25. [doi: 10.1002/pds.3533] [Medline: 24130046]
6. Hedberg K, Bui L, Livingston C, Shields L, Van Otterloo J. Integrating public health and health care strategies to address the opioid epidemic: the Oregon Health Authority's opioid initiative. J Public Health Manag Pract 2019;25(3):214-220. [doi: 10.1097/PHH.0000000000000849] [Medline: 30048336]
7. Marshall BDL, Yedinak JL, Goyer J, Green TC, Koziol JA, Alexander-Scott N. Development of a statewide, publicly accessible drug overdose surveillance and information system. Am J Public Health 2017 Nov;107(11):1760-1763. [doi: 10.2105/AJPH.2017.304007] [Medline: 28933938]
8. National Vital Statistics System. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/nchs/nvss/ [accessed 2020-03-18]
9. Stone AB, Jones MR, Rao N, Urman RD. A dashboard for monitoring opioid-related adverse drug events following surgery using a national administrative database. Am J Med Qual 2019 Jun 25;34(1):45-52. [doi: 10.1177/1062860618782646] [Medline: 29938518]
10. About the Survey. National Survey on Drug Use and Health. URL: https://nsduhweb.rti.org/respweb/about_nsduh.html [accessed 2020-03-18]

11. Treatment Episode Data Set. Substance Abuse and Mental Health Services Administration. URL: https://www.samhsa.gov/data/data-we-collect/teds-treatment-episode-data-set [accessed 2020-03-18]
12. Monitoring the future. URL: http://www.monitoringthefuture.org/ [accessed 2020-03-18]
13. Jiang Y, McDonald JV, Koziol J, McCormick M, Viner-Brown S, Alexander-Scott N. Can emergency department, hospital discharge, and death data be used to monitor burden of drug overdose in Rhode Island? J Public Health Manag Pract 2017;23(5):499-506. [doi: 10.1097/PHH.0000000000000514] [Medline: 28009694]
14. Strand MA, Eukel H. A primary prevention approach to the opioid epidemic. Am J Public Health 2019 Jun;109(6):861-863. [doi: 10.2105/ajph.2019.305060]
15. Davis CS, Green TC, Zaller ND. Addressing the overdose epidemic requires timely access to data to guide interventions. Drug Alcohol Rev 2016 Jul 01;35(4):383-386. [doi: 10.1111/dar.12321] [Medline: 26382016]
16. Ising A, Proescholdbell S, Harmon KJ, Sachdeva N, Marshall SW, Waller AE. Use of syndromic surveillance data to monitor poisonings and drug overdoses in state and local public health agencies. Inj Prev 2016 Apr 04;22 Suppl 1(Suppl 1):i43-i49. [doi: 10.1136/injuryprev-2015-041821] [Medline: 27044495]
17. McDonald DC, Carlson K, Izrael D. Geographic variation in opioid prescribing in the U.S. J Pain 2012 Oct;13(10):988-996 [FREE Full text] [doi: 10.1016/j.jpain.2012.07.007] [Medline: 23031398]
18. Milone MC. Laboratory testing for prescription opioids. J Med Toxicol 2012 Dec 1;8(4):408-416 [FREE Full text] [doi: 10.1007/s13181-012-0274-7] [Medline: 23180358]
19. Tenore PL. Advanced urine toxicology testing. J Addict Dis 2010 Oct 24;29(4):436-448. [doi: 10.1080/10550887.2010.509277] [Medline: 20924879]
20. Goossen W, Langford LH. Exchanging care records using HL7 V3 care provision messages. J Am Med Inform Assoc 2014 Oct 01;21(e2):e363-e368 [FREE Full text] [doi: 10.1136/amiajnl-2013-002264] [Medline: 24748607]
21. Abhyankar S, Demner-Fushman D, McDonald CJ. Standardizing clinical laboratory data for secondary use. J Biomed Inform 2012 Aug;45(4):642-650 [FREE Full text] [doi: 10.1016/j.jbi.2012.04.012] [Medline: 22561944]
22. Vreeman DJ, McDonald CJ, Huff SM. LOINC®: a universal catalogue of individual clinical observations and uniform representation of enumerated collections. IJFIPM 2010;3(4):273. [doi: 10.1504/ijfipm.2010.040211]
23. Slavova S, Bunn T, Hargrove S, Corey T, Ingram V. Linking death certificates, postmortem toxicology, and prescription history data for better identification of populations at increased risk for drug intoxication deaths. Pharm Med 2017 Apr 21;31(3):155-165 [FREE Full text] [doi: 10.1007/s40290-017-0185-7]
24. Dowell D, Haegerich T, Chou R. CDC Guideline for Prescribing Opioids for Chronic Pain — United States, 2016. Centers for Disease Control and Prevention. 2016. URL: https://www.cdc.gov/mmwr/volumes/65/rr/rr6501e1.htm?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fmmwr%2Fvolumes%2F65%2Frr%2Frr6501e1er.htm#contribAff [accessed 2023-09-07]
25. Manchikanti L, Abdi S, Atluri S, Balog CC, Benyamin RM, Boswell MV, American Society of Interventional Pain Physicians. American Society of Interventional Pain Physicians (ASIPP) guidelines for responsible opioid prescribing in chronic non-cancer pain: Part 2--guidance. Pain Physician 2012 Jul;15(3 Suppl):S67-116 [FREE Full text] [Medline: 22786449]
26. Lowder EM, Ray BR, Huynh P, Ballew A, Watson DP. Identifying unreported opioid deaths through toxicology data and vital records linkage: case study in Marion County, Indiana, 2011–2016. Am J Public Health 2018 Dec;108(12):1682-1687. [doi: 10.2105/ajph.2018.304683]
27. McClure FL, Niles JK, Kaufman HW, Gudin J. Concurrent use of opioids and benzodiazepines: evaluation of prescription drug monitoring by a United States laboratory. J Addict Med 2017;11(6):420-426 [FREE Full text] [doi: 10.1097/ADM.0000000000000354] [Medline: 28953504]
28. Phalen P, Ray B, Watson DP, Huynh P, Greene MS. Fentanyl related overdose in Indianapolis: estimating trends using multilevel Bayesian models. Addict Behav 2018 Nov;86:4-10 [FREE Full text] [doi: 10.1016/j.addbeh.2018.03.010] [Medline: 29631798]
29. Soderstrom CA, Dischinger PC, Kerns TJ, Kufera JA, Mitchell KA, Scalea TM. Epidemic increases in cocaine and opiate use by trauma center patients: documentation with a large clinical toxicology database. J Trauma 2001 Sep;51(3):557-564. [doi: 10.1097/00005373-200109000-00024] [Medline: 11535910]
30. Soper R, Appajosyula S, Deximo C. Decline in buprenorphine/naloxone prescriptions in a state Medicaid population following formulary conversion from suboxone to bunavail. Adv Ther 2018 Apr 6;35(4):457-466. [doi: 10.1007/s12325-018-0696-y] [Medline: 29623562]
31. ESSENCE Site. URL: https://essence.syndromicsurveillance.org/ [accessed 2020-03-18]
32. Lovegrove MC, Dowell D, Geller AI, Goring SK, Rose KO, Weidle NJ, et al. US Emergency Department visits for acute harms from prescription opioid use, 2016–2017. Am J Public Health 2019 May;109(5):784-791. [doi: 10.2105/ajph.2019.305007]
33. Zhou H, Burkom H, Winston CA, Dey A, Ajani U. Practical comparison of aberration detection algorithms for biosurveillance systems. J Biomed Inform 2015 Oct;57:446-455 [FREE Full text] [doi: 10.1016/j.jbi.2015.08.023] [Medline: 26334478]
34. Dodson ZM, Enki Yoo E, Martin-Gill C, Roth R. Spatial methods to enhance public health surveillance and resource deployment in the opioid epidemic. Am J Public Health 2018 Sep;108(9):1191-1196. [doi: 10.2105/ajph.2018.304524]

Abbreviations

ESSENCE: Electronic Surveillance System for the Early Notification of Community-Based Epidemics
HL7: Health Level 7
LOINC: Logical Observation Identifiers Names and Codes

Edited by E Mensah; submitted 14.10.20; peer-reviewed by J Russell; comments to author 08.12.20; accepted 11.01.23; published 28.09.23.

Please cite as:
Schleyer T, Robinson B, Parmar S, Janowiak D, Gibson PJ, Spangler V
Toxicology Test Results for Public Health Surveillance of the Opioid Epidemic: Retrospective Analysis
Online J Public Health Inform 2023;15:e50936
URL: https://ojphi.jmir.org/2023/1/e50936
doi:10.2196/50936
PMID:

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org