

People to policy: The promise and challenges of big data for India

Anthony Vipin Das^{1,2}

Big data holds great promise to help unravel insights to bridge the gap in human understanding. There has to be an emphasis on the quality of the data points being collected to ensure meaningful analysis. India has made significant strides to lay down a strong framework through the National Digital Health Blueprint and the National Health Stack for the future. There is a need to focus on the first important step of collection of a “good quality” data point through the implementation of electronic medical records by the health care providers. In India, 60 million individuals move below the poverty line every year because of the expenses related to unforeseen illness that adversely affects the individual’s welfare and the nation’s economic growth. With an out-of-pocket expense rate currently at 70% and the government’s health budget at a mere 1.3% of its GDP (gross domestic product), data-driven decisions are the need of the hour for policy making and to ensure equitable, efficient, and excellent delivery of health care. There is a huge potential to harness the power of big data to generate insights to address the four big challenges of health care in India – availability, accessibility, affordability, and acceptability.

Key words: Big data, electronic medical records, health care policy, India

Access this article online

Website:

www.ijo.in

DOI:

10.4103/ijo.IJO_1045_21

Quick Response Code:



As Good as the Data Point

Data is the new oil, or is it? The world today is currently navigating through the Fourth Industrial Revolution where the barrier between man and machine is dissolving.^[1] Data are being generated exponentially on who we were, who we are, and where we need to be. A data point is a discrete unit of information. It is generated from various sources such as the internet, financial sector, health care, mobile data, and others.^[2] The magnitude of data produced by humans every day is 2.5 quintillion bytes; about 1.72 MB of data is created every second by every person, and an astonishing 90% of the world’s data has been created in the past 2 years alone.^[3] *Big data* is a term that describes a large volume of data that may be analyzed computationally to reveal trends, associations, and patterns in the area of interest. There are specific attributes to big data that are more popularly known as the 5 “V”s and are defined as volume, velocity, veracity, variety, and value [Fig. 1].^[4] *Volume* is characterized by the sheer number of records or transactions being generated with time. *Velocity* is characterized by real-time creation or streaming of data at a constant rate. *Veracity* is characterized by trustworthiness and authenticity of the data being generated. *Variety* is characterized by structured, unstructured, and multimodal data sets. *Value* is finally dependent on all of the above attributes and provides insights to drive growth and impact. Health care data generated in ophthalmology encompasses all the five specific attributes of big data. The sheer volume of the number of patients examined,

the velocity of provision of eye care services both outpatient and surgical, the standardized recording of visual outcomes, intensive generation of different data points pertaining to text, images, and video make ophthalmology an area with high potential to generate value through big data. However, it is not just enough to be able to generate data; the mandate should be to generate “good quality” data in health care.

As Good as Our Weakest Link

It is important to define data quality to ensure value in its analysis. The five characteristics of data quality are accuracy, completeness, consistency, uniqueness, and timeliness.^[5] *Data Accuracy* ensures that the information that is being inputted into the database is correct and acceptable for future analysis. *Completeness* ensures a comprehensive collection of the information. *Consistency* ensures the uniformity in which the data are generated either in structured, semistructured, or unstructured formats. *Uniqueness* ensures whether the data captured are relevant for analysis. *Timeliness* ensures that the data are up-to-date and can be used for real-time reporting. In just 8 years, the estimated worldwide digital health care data grew from 50 petabytes in 2012 to 2,500 petabytes of data in 2020.^[6] Health care information system have evolved as an institutional backbone running its entire operations and clinical processes.^[7] Health care data are unique in the fact that they are complex and relational and comprise various formats including text, images, and media. The data quality in health care is of paramount importance to create best practice patterns, ensure

¹Department of eyeSmart EMR and AEye, L V Prasad Eye Institute, Hyderabad, Telangana, ²Indian Health Outcomes, Public Health and Economics Research Center, L V Prasad Eye Institute, Hyderabad, Telangana, India

Correspondence to: Dr. Anthony V Das, Department of eyeSmart EMR and AEye, L V Prasad Eye Institute, Hyderabad, Telangana - 500 034, India. E-mail: vipin@lvpei.org

Received: 31-Jul-2021

Revision: 16-Aug-2021

Accepted: 24-Aug-2021

Published: 29-Oct-2021

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: WKHLRPMedknow_reprints@wolterskluwer.com

Cite this article as: Das AV. People to policy: The promise and challenges of big data for India. Indian J Ophthalmol 2021;69:3052-7.

continuous quality improvement, and deliver personalized medicine to patients. User input variance is an important factor to be considered in health care data, and the use of structured clinical examination forms using electronic medical records (EMR) can greatly enhance the quality of the data set.^[8] Outcomes in health care are varied across the different disciplines that range from an immediate improvement in visual outcome post-cataract surgery compared with a prolonged rehabilitation and recovery in an orthopedic surgery. It is important to understand that documentation of health care outcomes is a function of time, and not all disciplines are similar. Research is a snapshot in time and evolves with the growing experience of delivering patient care and reporting the same. Data visualization tools allow the creation of real-time data pipelines from EMR to capture trends in disease burden, disease progression, and treatment outcomes that present the most current insights beyond the duration of the study.^[9,10] The creation of “continuous healthcare data pipelines” from EMR is a reality and must become the norm to benefit from the latest insights from the analysis of big data sets. Does size matter? Calculation of the sample size in the early stages of a study is a crucial step that affects the efficiency of the research being performed.^[11] The H-K-M-B-T pipeline [Fig. 2] aims to put things in perspective. Sample sizes that have been in

the range of hundreds and thousands are evolving into the millions today.^[12] Notable verticals such as finance and social media around the world have moved into analyzing billions and trillions of data points to generate meaningful insights to customize service delivery. Health care, unfortunately, is lagging behind, and we must explore if the revelation of new insights are different in the realm of the million, billion, and trillion data points. This is a path that must be taken into the unknown with each new discovery adding to the existing body of knowledge. There is a huge opportunity to bridge the gap in human understanding about various diseases. The ability of discovering patterns in big data that are beyond human comprehension poses the most exciting aspect of research using these large data sets. These unseen patterns can help in clinical decision support, disease surveillance, and population health management.^[13] The common ending for big data or traditional sampling-based inference is that as the sample size grows larger, the insights get stronger, and the longer the real-time follow-up, the results are closer to clinical significance and relevance. Big data science is an important inevitable complementary to evidence-based medicine. There is a need to generate continuous, consistent, and connected big data sets in health care to enable automation of insights using technology tools. Generation of a “good quality” data point is the first important step in our quest for new knowledge.

Electronic Medical Records: Boon or Bane

There is an increasing trend over a couple of decades among nations around the world to digitize health care processes to improve the quality of delivery and optimize the costs.^[14] The global health observatory data reported steady growth in the adoption of EMR over the past 15 years and 46% global increase in the past 5 years. More than 50% of the upper-middle- and high-income countries have adopted national EMR systems. The adoption of EMR is lower in the lower-middle- (35%) and low-income (15%) countries. The most frequently cited barriers to going digital were lack of infrastructure, capacity, funding, and legal frameworks.^[15] There is reasonable progress among 56% of the World Health Organization member countries in implementing legislation governing the EMR systems; however, a large proportion of countries lack appropriate legislation. At the end of 2016, the Healthcare Information and Management Systems Society, Asia-Pacific analysis adoption model of EMR reported that only 2.6% of hospitals achieved a Level 6 certification in India, which required physician documentation (templates), full Clinical Decision Support System (CDSS), and a closed-loop medication administration. Approximately 1.5% of the hospitals in India are measured at Level 7 of functioning, which involves complete EMR, Data

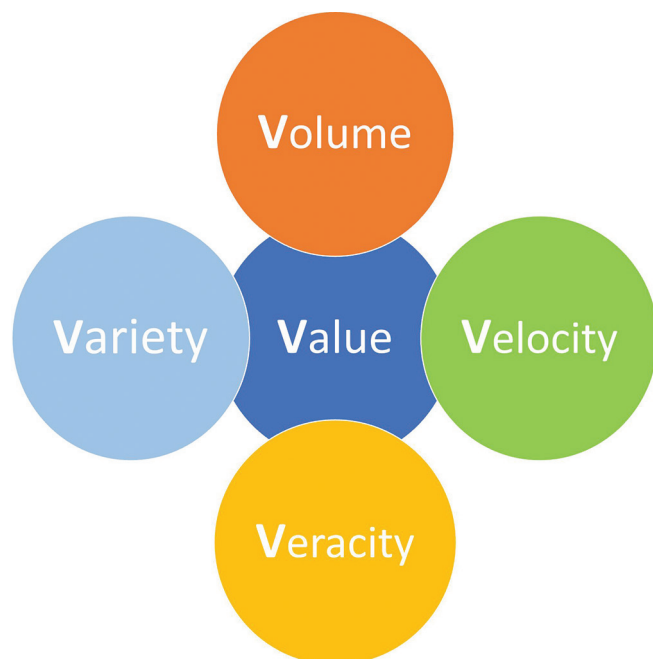


Figure 1: The 5 “V”s of big data



Figure 2: The evolution of the sample size in health care research

Analytics to improve care at the end of 2020. Currently, a larger proportion of hospitals (31%) are measured at Level 3 certification, which include Nursing/Clinical Documentation, CDSS (error checking), and PACS (picture archiving and communication system) available outside radiology.^[16] In April 2013, the Ministry of Health and Family Welfare, Government of India, had launched their recommendations of EMR standards in India and updated them in 2016.^[17] The activities in EMR adoption include creation of the basic Information and Communication Technology infrastructure, following the prescribed standards for development to ensure interoperability and adhering to the national policy and regulations.

The Data of the Nation

The data consciousness of India as a nation has evolved significantly over the past decade. The Open Government Data Platform, India, or data.gov.in supports the vision of Open Data initiative of the Government of India.^[18] The portal is a single-point access to documents, data sets, tools, services, and applications published by the ministries, departments, and organizations of the Government of India and was launched in October 2012. The initiative is in compliance with the National Data Sharing and Accessibility Policy of India, Gazette notified in March 2012.^[19] The Department of Health and Family Welfare has more than 301,054 resources, 511 catalogs, and 30,534 application programming interfaces (APIs). Some of the data sets pertaining to ophthalmology on the portal include reports of cataract operations performed, a pilot survey on causes of blindness, school screening programs, prophylaxis against blindness due to vitamin A deficiency, number of eye donations, states/union territories information of eye mishaps, annual record of eye surgeries performed in National Institutes of Health, and disabled population among main workers, marginal workers and nonworkers. The National Health Systems Resources Centre publishes the HMIS (Health Management Information System) Data Analysis that is based on the data uploaded to the National Web portal (<https://nhsrcindia.org/>) of all the states and districts of India. The health care data are analyzed on an annual basis, and the latest available data are 2015–2016, which is postdated.^[20] Here lies an opportunity to be able to collate information from EMR related to eye disease burden and outcomes both from the

public and private sectors. As a nation, India has taken the step forward to provide the data in a transparent manner to its citizens through the data.gov.in platform. There is a potential to harness deidentified data in real time from the health care providers from their EMR systems in a standard format from across the country to guide policy.

National Digital Health Blueprint

As an effort to ensure a holistic pursuit of the National Health Policy (NHP) 2017 and the Sustainable Development Goals related to health, the National Digital Health Blueprint (NDHB) was launched in 2019.^[21] Amid the current challenges with no interoperability of health record data of patient between service providers in the public and private sectors, no complete data for policy analysis and evidence-based interventions, the NDHB was launched after surveying the global best practices in the adoption of digital technologies. The key components of the NDHB include a set of architectural principles, a federated architecture, a five-layered system of architectural building blocks, Unique Health Identifier, data privacy and consent management, national portability, EMR, applicable standards and regulations, and health analytics and channels for the citizens to engage with the government such as the National Digital Health portal and the MyHealth App. The NHP also has prescribed goals of establishing disease registries for all diseases of public importance and, more important, a Federated National Health Information Architecture to link systems across the public and private health providers at the state and national levels, including the modern medicine systems, AYUSH (Ayurveda, Yoga and Naturopathy, Unani, Siddha, and Homeopathy) systems consistent with Metadata and Data Standards, and EMR. The National Digital Health Mission was established to drive the implementation of the blueprint and to promote and facilitate the evolution of the National Digital Health Ecosystem. Among the 35 building blocks identified in the NDHB, the key aspects include identification, citizen to be in control, service access/delivery, and interoperability. The NDHB recommends adopting the principle of “Think Big, Start Small, Scale Fast,” which is in line with the most important link in the entire digital health ecosystem – a “good quality” data point when generated exponentially achieves the intended goal of informed policy decisions for the people.

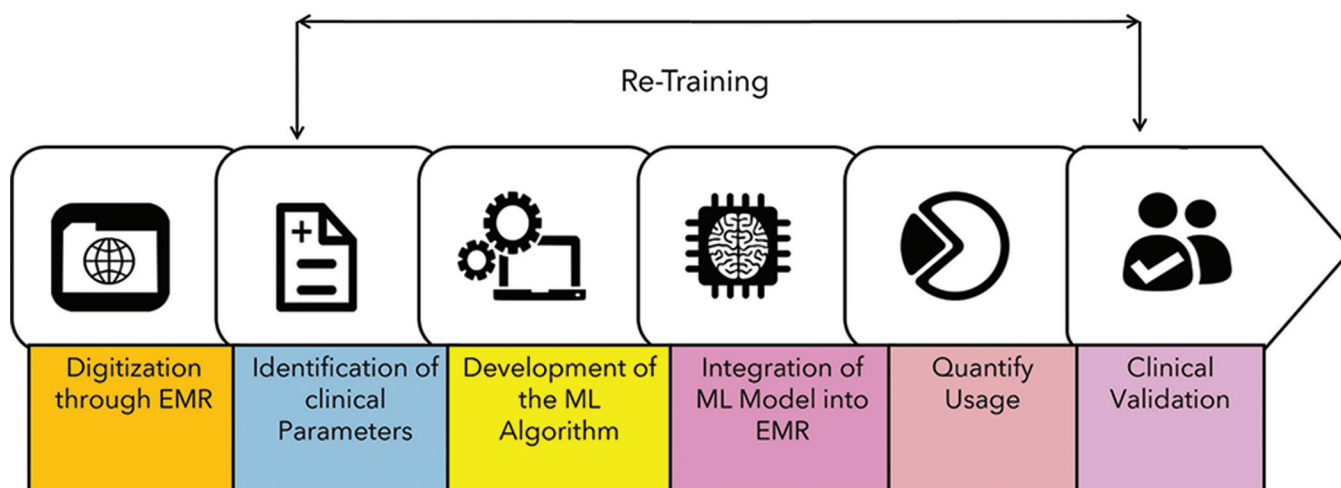


Figure 3: The AEye pipeline

The AEye Pipeline

The application of data science to big data sets in ophthalmology can be summarized in *The AEye Pipeline*.^[22] The pipeline comprehensively describes seven steps in the application of technology on clinical data sets for the generation of newer insights to the care providers. The steps are detailed in Fig. 3. Step 1 details the *Digitization of data through EMR*, which is the primary step for any kind of research based on big data sets. The IRIS (Intelligent Research in Sight) Registry is the largest national clinical specialty data registry with nearly 50 million patient visits and 14 million unique patients in the United States.^[23] The Système national d’information inter-régime de l’assurance maladie (SNIIR-AM) database covers nearly 65 million people and consists of medical and administrative data from 1,546 French private or public health care facilities catering to the general population.^[24] The National Ophthalmology Database from the United Kingdom was initiated by the Royal College of Ophthalmologists to collate anonymized clinical data from contributing hospitals from the National Health Service (NHS).^[25] The UK DR EMR (United Kingdom Diabetic Retinopathy Electronic Medical Record) Users Group has published their comprehensive experience of analyzing big data in patients with diabetic retinopathy from 20 UK hospital eye services on the real-world prevalence and progression data,^[26] impact of cataract surgery on diabetic macular edema,^[27] progression of diabetic retinopathy,^[28] and the impact of deprivation on the presentation of diabetic eye disease at hospitals services.^[29] The eyeSmart EMR database implemented in a multitier ophthalmology network in India has reported an 8-year experience describing 4.7 million patient visits and 2.2 million unique patients seen on EMR.^[8] The Ophthathome™ system is an integrated knowledge base of ophthalmic disease and has collected data of 0.5 million unique patients in India.^[30] The implementation of EMR systems in large ophthalmology institutions^[31] and the use of big data to understand endophthalmitis prophylaxis have been described in India.^[32] It is evident that the digitization of clinical data helps generate large data sets for further analysis. Step 2 details the *Identification of the clinical parameters* and the benefits with structured data sets and the challenges with unstructured ones. It is important to structure the input of clinical information in a standardized format as far as possible to ensure uniformity of the data for analysis. The use of drop-down menus, multiple selection, and standard forms enables the user to input the data in a repeatable manner. The use of free textboxes in EMR contributes to unstructured data and missing information that will lead to challenges during analysis.^[33] The use of deep learning techniques in natural language processing^[34] and finite state modeling to extract information from unstructured data entered in free textboxes^[35] can lead to increased information integrity. Step 3 is the *Development of machine learning (ML) algorithms* and the application of data science techniques on big data sets. Research using ML models such as linear regression, logistic regression, support vector machine, classification and regression trees, ensemble methods, and artificial neural network has the potential to improve ocular disease diagnosis, disease progression, and risk assessment.^[36] Deep learning algorithm focuses on three major aspects of classification, segmentation, and prediction using the static images generated from an ocular diagnostic device such as fundus photography, optical coherence tomography, visual

fields, and others.^[37] Step 4 details the *Integration of the ML model into the EMR* to complete the loop at the point of care. While a significant amount of research is advanced in the development of algorithms, it is important to integrate the insight into the EMR for efficient care of the patient. The entire purpose is to enhance the provision of care by providing the edge for superior clinical decision making by the care provider, which otherwise would not have been possible.^[38,39] Step 5 details *Quantification of the usage* of the insights provided by the machine. This is a crucial dynamic at the intersection of man and machine that evolves with time. How will a clinician with decades of experience trust an Artificial Intelligence (AI) insight? How can trust in the AI system be optimized to improve decision making? Interpersonal trust is a human belief based on three main dimensions of benevolence, integrity, and ability.^[40] The reputation of the system, the integrity of the insight, and the user perception of the ability of the AI all influence the reliance of a clinician on the technology.^[41] Step 6 is *Clinical validation* of the algorithm or the software as a Medical Device (SaMD) in a real-world setting. Outcomes in health care are influenced by many patient-related factors that can or cannot be quantifiable. Hence, tracking the performance of the algorithm over a period of time is important to build trust among clinicians using this technology. Clinical evaluation can either be a *Valid clinical association* that generates evidence to demonstrate the association between SaMD output and SaMD-targeted clinical condition, or an *Analytical validation* that generated evidence that the SaMD correctly processes input data to generate accurate, reliable, and precise output data, or *Clinical validation* that generates evidence that the SaMD’s accurate, reliable, and precise data output achieves its intended purpose in its target population in the context of clinical care.^[42] There is a need for assurance for safety and effectiveness of any SaMD by ensuring high data quality for training, algorithm correctness (verification), performance testing (validation), and generalizability (addressing bias) and reliable interpretability. Step 7 completes the cycle by incorporating *Retraining* or continuous learning in real time. The SaMD must be capable of capturing real-world performance data to understand the user interactions with the SaMD and must be able to monitor its technical and analytical performance to evolve into a more efficient and reliable tool for the clinicians. This continuous feedback will connect back to Step 2 to include or exclude relevant parameters as an input into the ML model to further refine the SaMD. We need to avoid any unintended consequences to the patient due to a wrong insight that can be precipitated due to a break in any of the

Table 1: The way forward with big data for India: People to policy

| | |
|-----------|---|
| Integrate | Federated databases (EMR/data registries) |
| | Diagnostic centers/biobanks |
| | Epidemiology consortia |
| Insights | Citizen centric |
| | Wellness centric |
| | Informed decision making |
| Impact | Identify unmet needs |
| | Track metrics |
| | Monitor feedback |

EMR=electronic medical records

above seven steps of the AEye pipeline. The key to improving health care delivery is digitization through EMR, implementing SaMD and continuous learning.

People to Policy

There is nothing more powerful than data generated from the people informing policy decisions that affect them positively. The implementation of EMR today has enabled the study of disease presentations,^[43-46] disease prognosis,^[47] disease outcomes,^[48,49] environment factors,^[50,51] and prediction of outcomes^[52] in big data sets. Although the intent is to utilize technology solutions to increase the efficiency of health care delivery, we must act in the intended direction in India. The three “I” concept of *Integrate*, *Insights*, and *Impact* can help bring the mandate together [Table 1]. There is a need to *Integrate* the various sources of data being generated from EMR or data registries, health insurance and administrative databases, diagnostic centers, biobanks, epidemiology consortia, and crowdsourced data through the proposed NHS by the NITI Aayog.^[53] The *Insights* generated from the integration of the data will help identify areas of unmet need and also track the progress of interventions and guide policy decisions. The greatest impact of data-driven informed decisions for the first time ever will be the shift toward citizen- and wellness-centric policy making in India. The integration of data and insights generated will catalyze *Impact* for the people. There is a need for the government to bring together health care providers who have currently digitized their data to show the potential of collating this standardized information to impact health care delivery for the population. We need to “Walk the Talk!” There is an amazing opportunity for us to consolidate the strides we have made so far in our nation in multiple segments and bring them together to show a functional model of how digitized data can transform health care in India. It is the responsibility of the health care providers to realize the potential for good and lead the way by taking the first step of adopting EMR.

Conclusion

Catering to the health care needs of a billion people poses unique challenges and opportunities in a nation like India. The provision of the basic right to health care is the moral responsibility of the government to its citizens. Although access to health information is crucial, the emphasis must be on the quality of the data being collected that will guide the decisions of our policy makers. We must and should adopt the best technology frameworks that are relevant to our geography without compromising the volume of care that is currently being delivered. With the changing landscape of both lifestyle and the trends of diseases over the years, it is imperative that we need to have the right information at the right time for the right decision for the right individual. The investment in taking the bold step of digital transformation for health care institutions in India will pay rich dividends for the nation and its citizens in the long run. The implementation of digital health is a key step in the direction to ensure the availability, accessibility, affordability, and acceptability of health care for the masses. There has never been a more opportune time such as this to lay down a strong foundation to enable the collection of “good quality” digital health data points to guide us in the future! The willingness to take this first step by every health

care provider is a crucial drop contributing to the vast ocean of digital health transformation of India, thereby enabling the delivery of equitable, efficient, and excellent health care to all those in need.

Financial support and sponsorship

This study was supported by the Hyderabad Eye Research Foundation, Telangana, India.

Conflicts of interest

There are no conflicts of interest.

References

1. Available from: https://en.wikipedia.org/wiki/Fourth_Industrial_Revolution. [Last accessed on 2021 Apr 11].
2. Lei J, Kong L. 2- Fundamentals of big data in radio astronomy. *Big Data in Astronomy*. Elsevier; 2020. p. 29-58.
3. Available from: <https://techjury.net/blog/how-much-data-is-created-every-day/#gref>. [Last accessed on 2021 Apr 11].
4. Ishwarappa, Anuradha J. A brief introduction on Big data 5Vs characteristics and hadoop technology. *Procedia Comput Sci* 2015;48:319-24.
5. Available from: <https://www.computerbusinessresearch.com/Home/database/five-characteristics-of-high-quality-information/>. [Last accessed on 2021 Apr 11].
6. Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. 2011. Available from: <https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation>.
7. Kaur K, Rani R. Managing data in healthcare information systems: Many models, one solution. *Computer* 2015;48:52-9.
8. Das AV, Kammari P, Vadapalli R, Basu S. Big data and the eyeSmart electronic medical record system-An 8-year experience from a three-tier eye care network in India. *Indian J Ophthalmol* 2020;68:427-32.
9. Das AV, Rath S, Naik MN, Ali MJ. The incidence of lacrimal drainage disorders across a tertiary eye care network: Customization of an indigenously developed electronic medical record system-eyeSmart. *Ophthalmic Plast Reconstr Surg* 2019;35:354-6.
10. Das AV, Kammari P, Guizie EB, Varada R, Singh G. Prevalence of ocular disorders in Liberia: A retrospective study using the eyeSmart electronic medical record system. *J Glob Health Rep* 2019;3:e2019033.
11. Das S, Mitra K, Mandal M. Sample size calculation: Basic principles. *Indian J Anaesth* 2016;60:652-6.
12. Donthineni PR, Kammari P, Shanbhag SS, Singh V, Das AV, Basu S. Incidence, demographics, types and risk factors of dry eye disease in India: Electronic medical records driven big data analytics report I. *Ocul Surf* 2019;17:250-6.
13. Liang Y, Kelemen A. Big data science and its applications in health and medical research: Challenges and opportunities. *J Biom Biostat* 2016;7:307. doi: 10.4172/2155-6180.1000307.
14. Health Information and Management Systems Society. *Electronic Health Records: A Global Perspective*. 2nd ed. Chicago (IL): Healthcare Information and Management Systems Society; 2010.
15. Available from: https://www.who.int/goe/publications/global_diffusion/en/. [Last accessed on 2021 Apr 11].
16. Sharma M, Aggarwal H. EHR adoption in India: Potential and the challenges. *Indian J Sci Technol* 2016;9:1-7.
17. Available from: https://www.nhp.gov.in/ehr_standards_mtl_mtl. [Last accessed on 2021 Apr 11].
18. Available from: <https://data.gov.in/> [Last accessed on 2021 Apr 12].

19. Available from: <https://dst.gov.in/national-data-sharing-and-accessibility-policy-0>. [Last accessed on 2021 Apr 12].
20. Available from: <http://www.nhs.uk/nhs.uk/hmis-data-analysis>. [Last accessed on 2021 Apr 12].
21. Available from: <https://main.mohfw.gov.in/newshighlights/final-report-national-digital-health-blueprint-ndhb>. [Last accessed on 2021 Apr 12].
22. Verkicharla PK, Das AV. Technology and myopia. *Community Eye Health* 2019;32:S9-10.
23. Parke I DW, Lum F, Rich WL. The IRIS® Registry: Purpose and perspectives. *Ophthalmology* 2017;114(Suppl 1):1-6.
24. Daien V, Korobelnik JF, Delcourt C, Cougnard-Gregoire A, Delyfer MN, Bron AM, *et al.* French medical-administrative database for epidemiology and safety in ophthalmology (EPISAFE): The EPISAFE collaboration program in cataract surgery. *Ophthalmic Res* 2017;58:67-73.
25. Jackson TL, Donachie PH, Sparrow JM, Johnston RL. United Kingdom National Ophthalmology Database Study of Vitreoretinal Surgery: Report 1; case mix, complications, and cataract. *Eye (Lond)* 2013;27:644-51.
26. Egan CA, Lee A, Zhu H, Crabb D, Tufail A, Johnston R; UK DR EMR Users Group. UK DR EMR Users Group: Report 1 - real world prevalence and progression data for diabetic retinopathy for 56 211 patients from 20 United Kingdom hospital eye services. *Invest Ophthalmol Vis Sci* 2015;56:1448.
27. Denniston AK, Chakravarthy U, Zhu H, Lee AY, Crabb DP, Tufail A, *et al.*; UK DR EMR Users Group. The UK Diabetic Retinopathy Electronic Medical Record (UK DR EMR) Users Group, Report 2: Real-world data for the impact of cataract surgery on diabetic macular oedema. *Br J Ophthalmol* 2017;101:1673-8.
28. Lee CS, Lee AY, Baughman D, Sim D, Akelere T, Brand C, *et al.*; UK DR EMR Users Group. The United Kingdom Diabetic Retinopathy Electronic Medical Record Users Group: Report 3: Baseline retinopathy and clinical features predict progression of diabetic retinopathy. *Am J Ophthalmol* 2017;180:64-71.
29. Denniston AK, Lee AY, Lee CS, Crabb DP, Bailey C, Lip PL, *et al.*; UK DR EMR Users Group. United Kingdom Diabetic Retinopathy Electronic Medical Record (UK DR EMR) Users Group: Report 4, real-world data on the impact of deprivation on the presentation of diabetic eye disease at hospital services. *Br J Ophthalmol* 2019;103:837-43.
30. Raj P, Tejwani S, Sudha D, Muthu Narayanan B, Thangapandi C, Das S, *et al.* Ophthatome™: An integrated knowledgebase of ophthalmic diseases for translating vision research into the clinic. *BMC Ophthalmol* 2020;20:442.
31. Scholl J, Syed-Abdul S, Ahmed LA. A case study of an EMR system at a large hospital in India: Challenges and strategies for successful adoption. *J Biomed Inform* 2011;44:958-67.
32. Haripriya A, Chang DF, Ravindran RD. Endophthalmitis reduction with intracameral moxifloxacin prophylaxis: Analysis of 600 000 surgeries. *Ophthalmology* 2017;124:768-75.
33. Bowman S. Impact of electronic health record systems on information integrity: Quality and safety implications. *Perspect Health Inf Manag* 2013;10:1c.
34. Ting DSW, Peng L, Varadarajan AV, Keane PA, Burlina PM, Chiang MF, *et al.* Deep learning in ophthalmology: The technical and clinical considerations. *Prog Retin Eye Res* 2019;72:100759.
35. Sai Prashanthi G, Deva A, Vadapalli R, Das AV. Automated categorization of systemic disease and duration from electronic medical record system data using finite-state machine modeling: Prospective validation study. *JMIR Form Res* 2020;4:e24490.
36. Lin WC, Chen JS, Chiang MF, Hribar MR. Applications of artificial intelligence to electronic health record data in ophthalmology. *Transl Vis Sci Technol* 2020;9:13.
37. Tong Y, Lu W, Yu Y, Shen Y. Application of machine learning in ophthalmic imaging modalities. *Eye Vis (Lond)* 2020;7:22.
38. Sagkriotis A, Chakravarthy U, Griner R, Doyle O, Wintermantel T, Clemens A. Application of machine learning methods to bridge the gap between non-interventional studies and randomized controlled trials in ophthalmic patients with neovascular age-related macular degeneration. *Contemp Clin Trials* 2021;104:106364.
39. Lin H, Long E, Ding X, Diao H, Chen Z, Liu R, *et al.* Prediction of myopia development among Chinese school-aged children using refraction data from electronic medical records: A retrospective, multicentre machine learning study. *PLoS Med* 2018;15:e1002674.
40. Mayer RC, Davis JH, Schoorman FD. An integrative model of organizational trust. *Acad Manag Rev* 1995;20:709-34.
41. Lee JD, See KA. Trust in automation: Designing for appropriate reliance. *Hum Factors* 2004;46:50-80.
42. Available from: <https://www.fda.gov/media/143303/download>. [Last accessed on 2021 Apr 16].
43. Kaliki S, Das AV. Ocular and periocular tumors in India: An EyeSmart electronic medical record analysis of 9633 cases from a referral center. *Middle East Afr J Ophthalmol* 2021;27:199-203.
44. Das AV, Donthineni PR, Sai Prashanthi G, Basu S. Allergic eye disease in children and adolescents seeking eye care in India: Electronic medical records driven big data analytics report II. *Ocul Surf* 2019;17:683-9.
45. Das AV, Podila S, Prashanthi GS, Basu S. Clinical profile of pterygium in patients seeking eye care in India: Electronic medical records-driven big data analytics report III. *Int Ophthalmol* 2020;40:1553-63.
46. Das AV, Basu S. Epidemic keratoconjunctivitis in India: Electronic medical records-driven big data analytics report IV. *Br J Ophthalmol* 2020;bjophthalmol-2020-317330.
47. Das AV, Basu S. Indications and prognosis for keratoplasty in eyes with severe visual impairment and blindness due to corneal disease in India. *Br J Ophthalmol* 2021;105:17-21.
48. Khanna RC, Rath VM, Guizie E, Singh G, Nishant K, Sandhu S, *et al.* Factors associated with visual outcomes after cataract surgery: A cross-sectional or retrospective study in Liberia. *PLoS One* 2020;15:e0233118.
49. Donthineni PR, Das AV, Shanbhag SS, Basu S. Cataract surgery in dry eye disease: Visual outcomes and complications. *Front Med (Lausanne)* 2020;7:575834.
50. Das AV, Basu S. Epidemic keratoconjunctivitis in India: Trend analysis and implications for viral outbreaks. *Indian J Ophthalmol* 2020;68:732-6.
51. Garrigan H, Ifantides C, Prashanthi GS, Das AV. Biogeographical and altitudinal distribution of cataract: A nine-year experience using electronic medical record-driven big data analytics in India. *Ophthalmic Epidemiol* 2020;1-8. doi: 10.1080/09286586.2020.1849741.
52. Achiron A, Gur Z, Aviv U, Hilely A, Mimouni M, Karmona L, *et al.* Predicting refractive surgery outcome: Machine learning approach with big data. *J Refract Surg* 2017;33:592-7.
53. Available from: https://niti.gov.in/writereaddata/files/document_publication/NHS-Strategy-and-Approach-Documents-for-consultation.pdf. [Last accessed on 2021 Apr 17].