

## Using Big Data in eye research to answer important scientific questions

### O Big Data vai nos ajudar a responder as grandes questões em pesquisa?

Monica Alves<sup>1</sup>, Rosalia Antunes Forschini<sup>2</sup>, Paulo Schor<sup>3</sup>

1. Department of Ophthalmology and Otorhinolaryngology, Faculty of Medical Sciences, Universidade Estadual de Campinas, Campinas, SP, Brazil.

2. Department of Ophthalmology, Otorhinolaryngology and Head and Neck Surgery, Faculdade de Medicina, Universidade de São Paulo, Ribeirão Preto, SP, Brazil.

3. Department of Ophthalmology, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP, Brazil.

Some of us remember how to deal with instructions as USE, SKIP, GO TOP and GO BOTTOM, especially if that comes together with a 286 PC black screen, the late 80's memories and a couple of old 3.5 floppy disks named Dbase.

This simple and limited knowledge evolved to a higher level when tons of records (volume) merged with velocity (to acquire and process it), variety and veracity. The aforementioned characteristics (the four "Vs") define Big Data, which was catalyzed by the Internet phenomenon that started in 1990s.

The Big Data concept has captured the attention of researchers worldwide as it offers an endless amount of patterns and predictions, alarming as much as interesting. Platforms provided a plethora of data, such as patient demographics, and clinical, ancillary, and even genomic data, which can be helpful for conducting large-scale, low-cost healthcare analysis and treatment decision-making.

Contemporary democratization included access to the "Net," which brings together major, free databases and mathematical models (algorithms) that manipulate this information, allowing us to live in an era of artificial intelligence.

Contests that challenge young minds in search of the best convolutional algorithm (that retrofeeds) using

open access registers are pictured in movies, such as the "The Social Network" from 2010. Moreover, currently, being able to code is as important as being able to read and write.

We live in a "camera culture" society (named after a Massachusetts Institute of Technology media-lab group) where there are more cameras than eyes. Where else could we extract more features (volume, variety, and speed) than medical (veracity) images? Image-related subspecialties, such as radiology and ophthalmology, are a natural locus to decouple pictures and experiment with their colors, contrast, and spatial distribution.

Our major role as clinicians is to classify (veracity) such images based on clinical diagnosis and most importantly, to find real relevance in the subject.

Several examples of digital classifications are available in ophthalmology, such as diabetic retinopathy gradings. In such cases, we teach machines how to use human impressions and separate retinal image patterns to digitally classify them. Continuous feed of categorized medical pictures allows the machine to screen cases for suspected diseases, thus reducing healthcare costs and improving healthcare efficacy.

Surprisingly, an algorithm can take into account huge amounts of information that is not important to humans and can even separate retinas of men from those of woman, which is an impossible task for us<sup>(1)</sup>. Is this artificial intelligence? Will we be able to be tutored by machines?

This seems likely, and several correlations and conclusions are being drafted at this very moment, offering personalized music and movies. This may be the basis for personalized medicine that is currently being referred to as "Network Medicine." Data, tools and artificial

Submitted for publication: September 12, 2018  
Accepted for publication: October 1, 2018

**Funding:** This study received no specific financial support.

**Disclosure of potential conflicts of interest:** None of the authors have any potential conflicts of interest to disclose.

**Corresponding author:** Paulo Schor.

Rua Pedro de Toledo, 781 - São Paulo, SP - 04039-032 - Brasil  
E-mail: pschor@pobox.com

 This content is licensed under a Creative Commons Attribution 4.0 International License.

intelligence are widely available to create it; however, experienced ophthalmologists still are required in this process. People who know the method of implementing artificial intelligence (even while not being the fastest experts in its use) and have daily access to the users (persons, community, clinics, and hospitals) add value to the solution and can improve practical application.

Developments, such as the recent, Preclinical Alzheimer OCT-A findings, will still be described as traditional research (“I suspect that the foveal avascular zone is a clue to the disease”); however, unexpected conclusions may seem magical, scaring us<sup>(2)</sup>. But even scarier are the possibilities to manipulate algorithms and produce biased conclusions. As scientists and critical clinicians, it is our duty to recognize possibilities of data misuse and, to be aware of alternatives to avoid that corruption, as split learning ([https://www.media.mit.edu/publications/split-learning-for-health-distributed-deep-learning-without-](https://www.media.mit.edu/publications/split-learning-for-health-distributed-deep-learning-without-sharing-raw-patient-data/)

[sharing-raw-patient-data/](https://www.media.mit.edu/publications/split-learning-for-health-distributed-deep-learning-without-sharing-raw-patient-data/)). The world is getting more complex, and still much more interesting.

Companies such as “23andMe” have more than 5 million users who are genetically screened and experimented using gene-based dietary regimens. It is a unique opportunity to actively participate in the evolution of humanity with the application of the natural intelligence. Let us give it relevance!

## REFERENCES

1. O’Byrhim BE, Apte RS, Kung N, Coble D, Van Stavern GP. Association of preclinical alzheimer disease with optical coherence tomographic angiography findings. *JAMA Ophthalmol*. Published online August 23, 2018. doi:10.1001/jamaophthalmol.2018.3556
2. Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng* [Internet]. 2018 [cited 2018 jan 21];2:158-64. Available from <https://static.googleusercontent.com/media/research.google.com/pt-BR//pubs/archive/46425.pdf>