Sistemas probabilísticos de apoio à decisão usando bigdata:

Gabriel Magalhães N. Guimarães, TSA-SBA, MSc
gabriel@gabriel.med.br
t.me/gabrielmng

Conflitos de interesse

- Nenhum relacionado com a apresentação.

http://gabriel.med.br/

Sociedade Brasileira de Anestesiologia

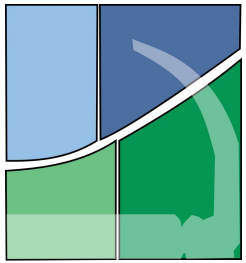
SOCIEDADE DE ANESTESIOLOGIA DO DISTRITO FEDERAL

sbis
sociedade brasileira de informática em saúde

SBED
SOCIEDADE BRASILEIRA PARA ESTUDO DA DOR

ABRAMEDE
ASSOCIAÇÃO BRASILEIRA DE MEDICINA DE EMERGÊNCIA

EBSERH
HOSPITAIS UNIVERSITÁRIOS FEDERAIS
MINISTÉRIO DA EDUCAÇÃO



ANESTESIOLOGIA
HUB

ICS
Instituto Castro e Santos

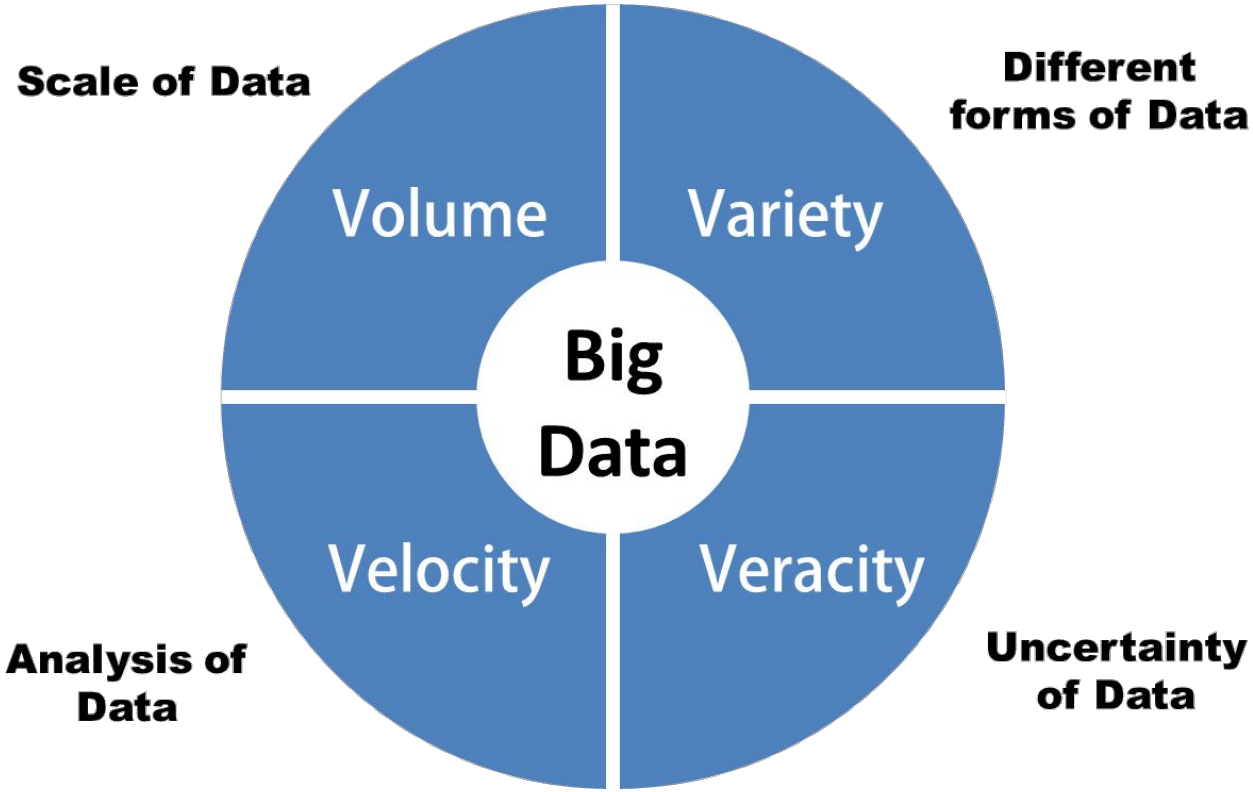
ANESTESIOLOGIA
USP

Universidade de Brasília

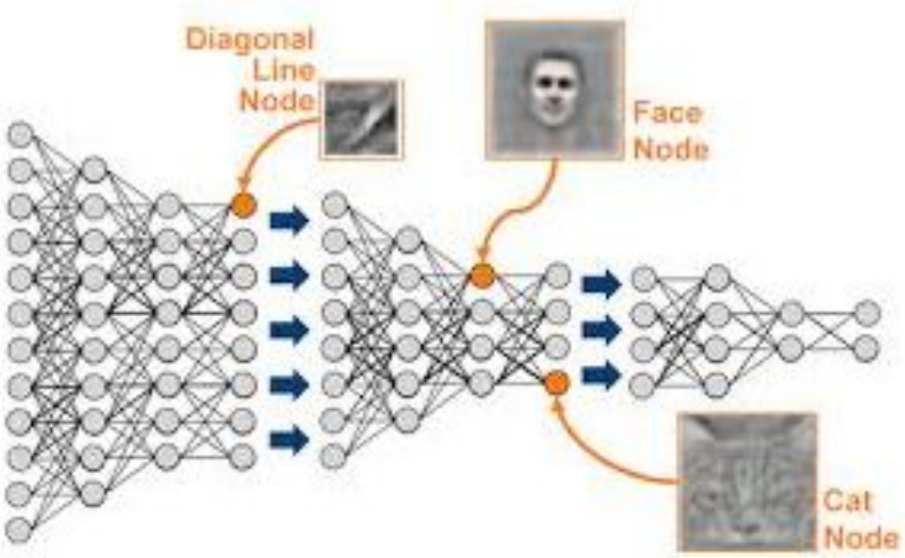
GRIFE: Modelagem de sistemas complexos

http://dgp.cnpq.br/dgp/espelhogrupo/4297092389120709

Big data?


Scale of Data
Different forms of Data
Volume
Variety
Big Data
Velocity
Veracity
Analysis of Data
Uncertainty of Data

Big data?



- Machine learning
- Processamento paralelo
- Técnicas inteligência artificial
- Redes neurais

N Engl J Med. 2016 September 29; 375(13): 1216–1219. doi:10.1056/NEJMp1606181.

Predicting the Future — Big Data, Machine Learning, and Clinical Medicine

Ziad Obermeyer, M.D., M.Phil. and
Department of Emergency Medicine, Harvard Medical School and Brigham and Women's Hospital, and the Department of Health Care Policy, Harvard Medical School, Boston

Ezekiel J. Emanuel, M.D., Ph.D.
Department of Medical Ethics and Health Policy, Perelman School of Medicine, and the Department of Health Care Management, the Wharton School, University of Pennsylvania, Philadelphia

N Engl J Med. 2016 September 29; 375(13): 1216–1219. doi:10.1056/NEJMp1606181.

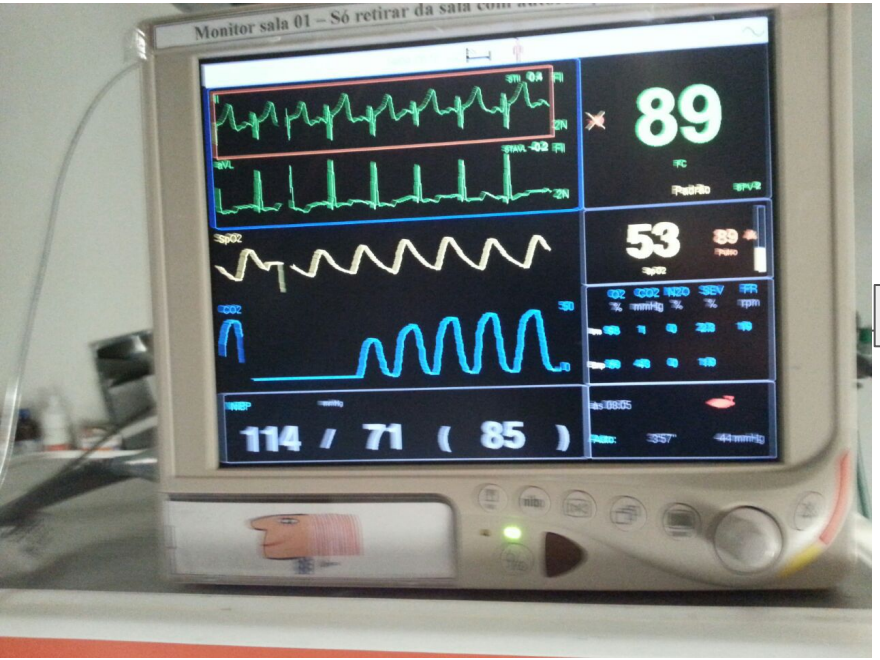
Predicting the Future — Big Data, Machine Learning, and Clinical Medicine

Ziad Obermeyer, M.D., M.Phil. and
Department of Emergency Medicine, Harvard Medical School and Brigham and Women's Hospital, and the Department of Health Care Policy, Harvard Medical School, Boston

Ezekiel J. Emanuel, M.D., Ph.D.
Department of Medical Ethics and Health Policy, Perelman School of Medicine, and the Department of Health Care Management, the Wharton School, University of Pennsylvania, Philadelphia

Second, machine learning will displace much of the work of radiologists and anatomical pathologists. These physicians focus largely on interpreting digitized images, which can easily be fed directly to algorithms instead. Massive imaging data sets, combined with recent advances in computer vision, will drive rapid improvements in performance, and machine accuracy will soon exceed that of humans. Indeed, radiology is already part-way there: algorithms can replace a second radiologist reading mammograms[5] and will soon exceed human accuracy. The patient-safety movement will increasingly advocate use of algorithms

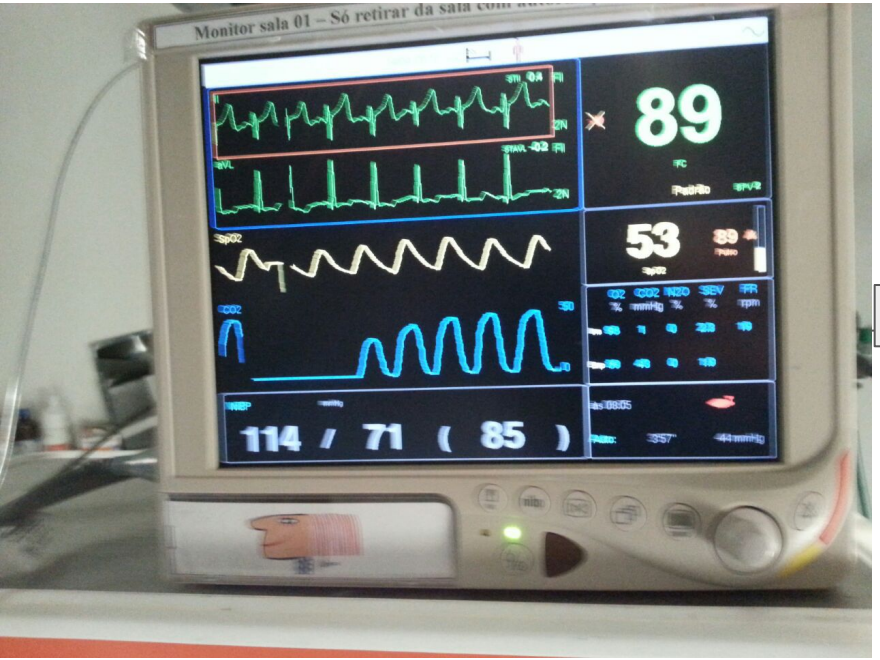
Um paciente, um instante



1 Unidade / min

FC
SpO2
EtCO2
PAS
PAD
PAM
+30

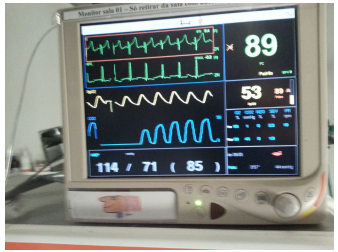
Um paciente, uma cirurgia curta



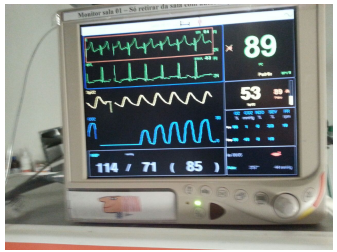
3h cirurgia = 180 linhas

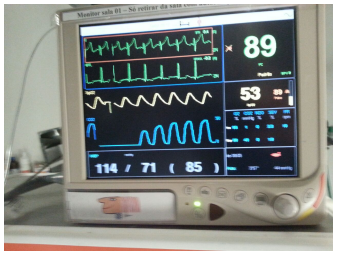
1 Unidade / min

FC
SpO2
EtCO2
PAS
PAD
PAM
+30



Um dia, uma sala cirúrgica





3 cirurgias = 540 linhas

1 Unidade / min

FC
SpO2
EtCO2
PAS
PAD
PAM
+30

Um dia, um centro cirúrgico 10 SO

5400 linhas

1 Unidade / min

FC
SpO2
EtCO2
PAS
PAD
PAM
+30

Um ano, um centro cirúrgico 10 SO

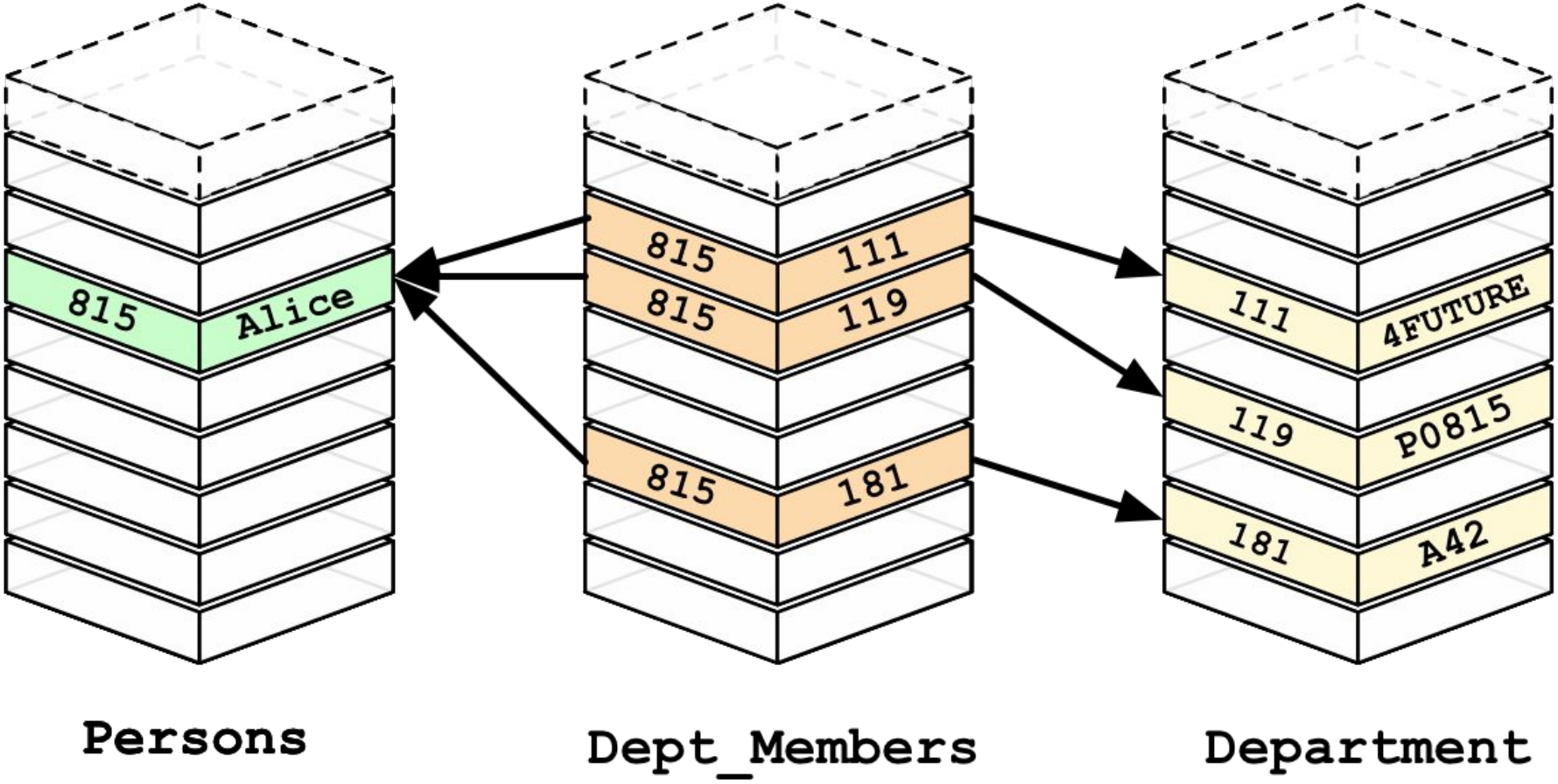
1.971.000 linhas

Um ano, dados multicentricos (50 hospitais)

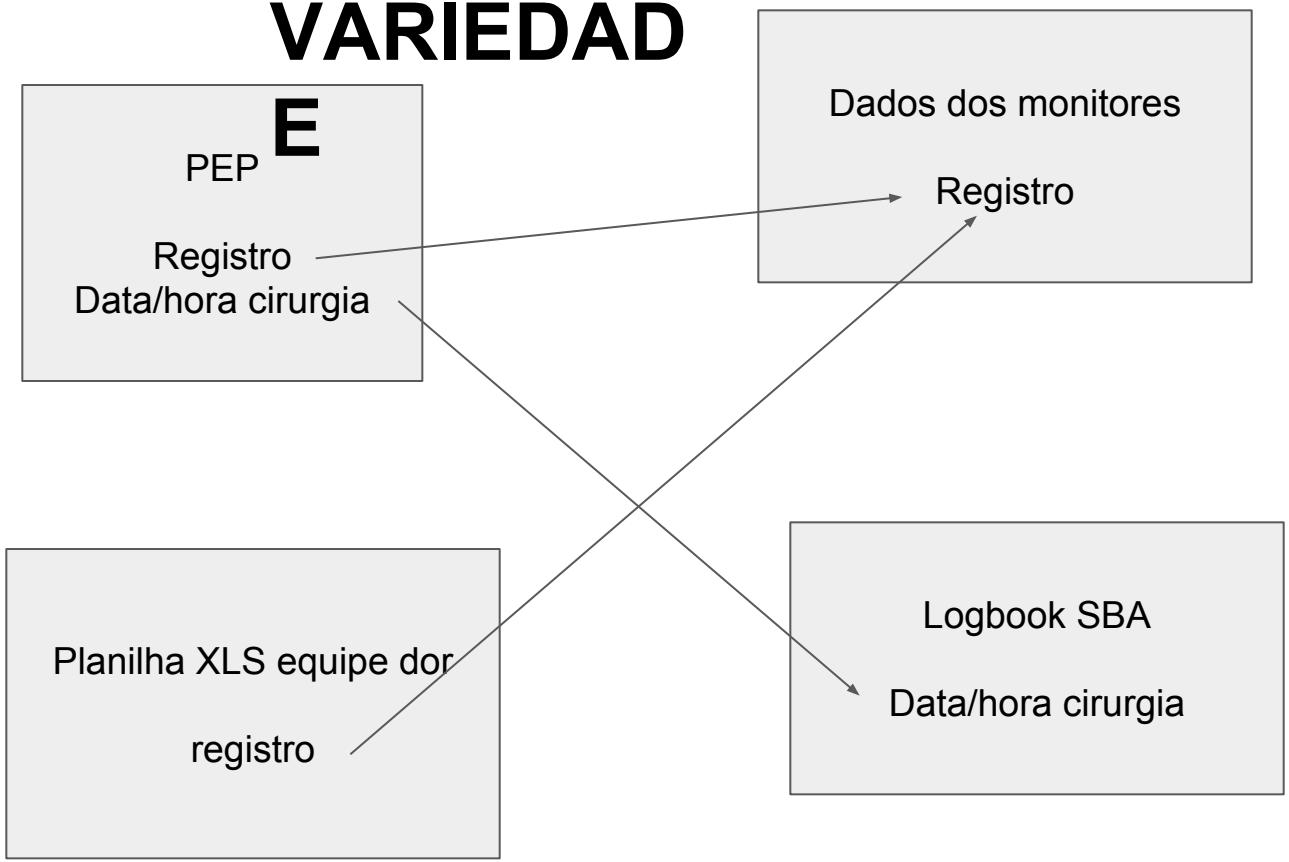
98.550.000 linhas

VOLUME

VELOCIDADE


815
Alice
815
111
815
119
815
181
111
4FUTURE
119
P0815
181
A42
Persons
Dept_Members
Department

VARIEDADE


PEP
Registro
Data/hora cirurgia
Dados dos monitores
Registro
Planilha XLS equipe dor
registro
Logbook SBA
Data/hora cirurgia

VERACIDADE

- Fenômeno da amortização;
- Dados incompletos;
- Dados inconsistentes;
  - ex: hora diferente bd diferente.
  - ex: variáveis subjetivas
- Múltiplos sinônimos / erros de grafia;
  - Variáveis categóricas.

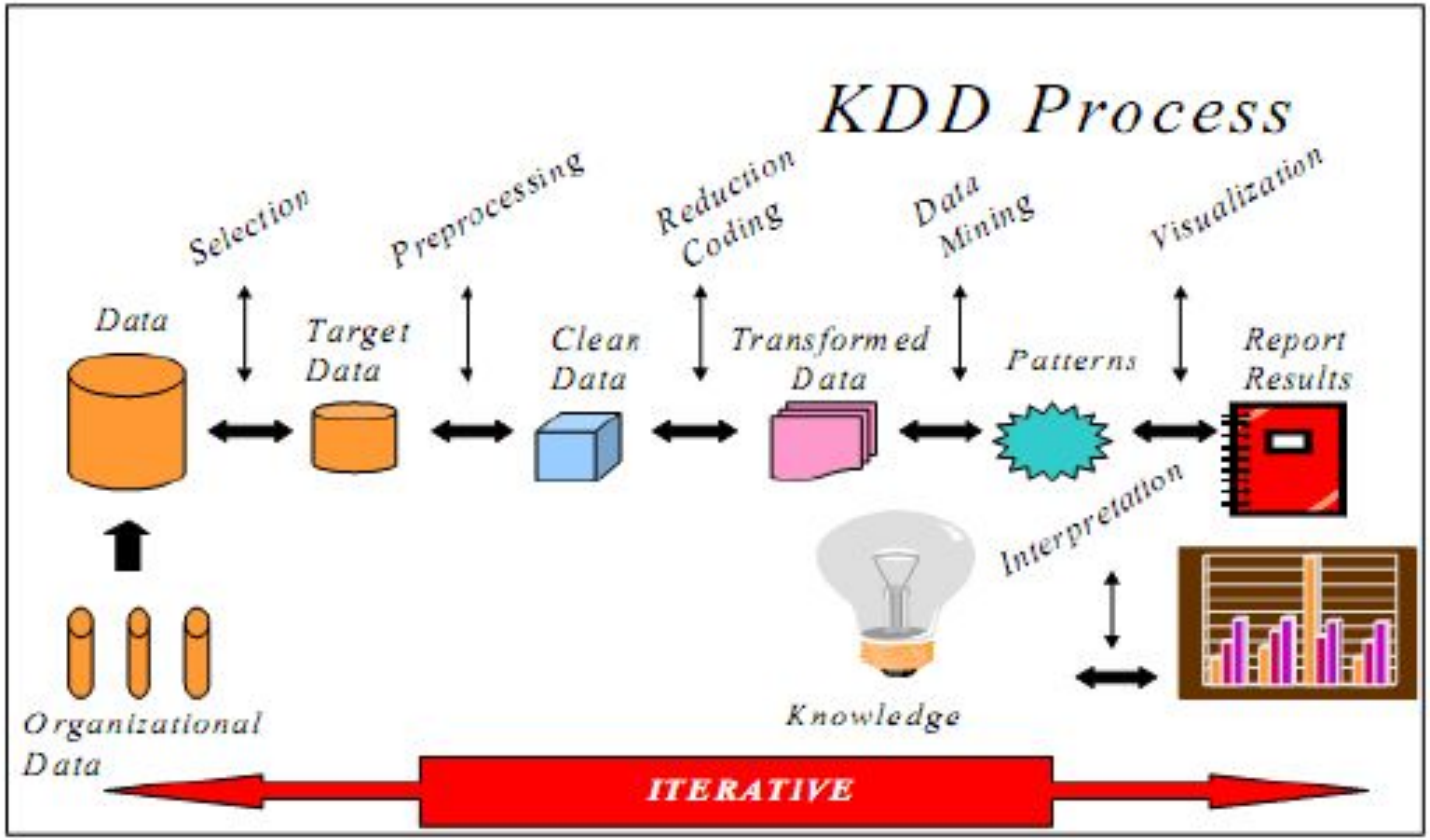
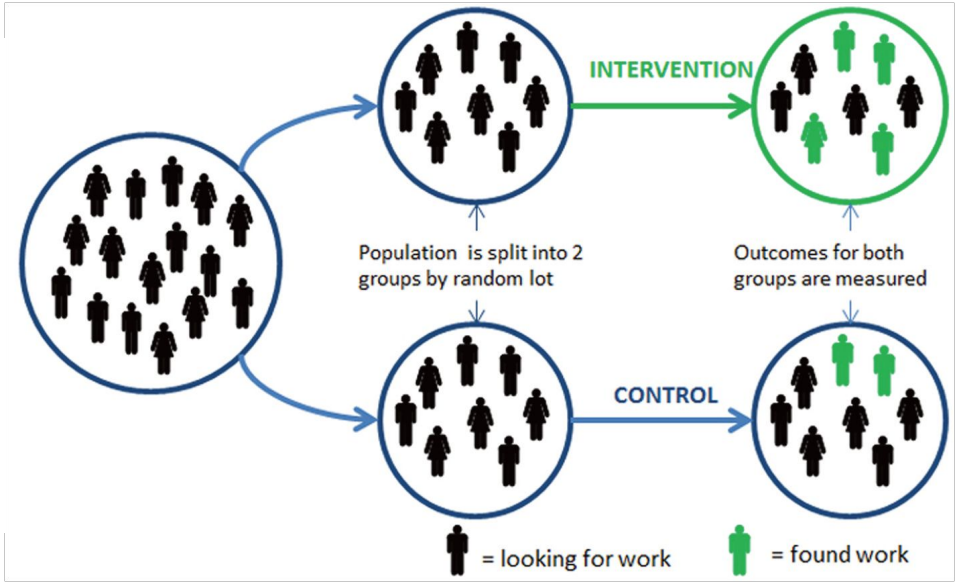


Figure 2: The KDD Process

Data mining VS Machine learning

- **Supervisionado vs não supervisionado;**
- **Objetivo específico (ex: pesquisa) vs modelo dinâmico;**
- **Pré-processamento manual vs automático;**

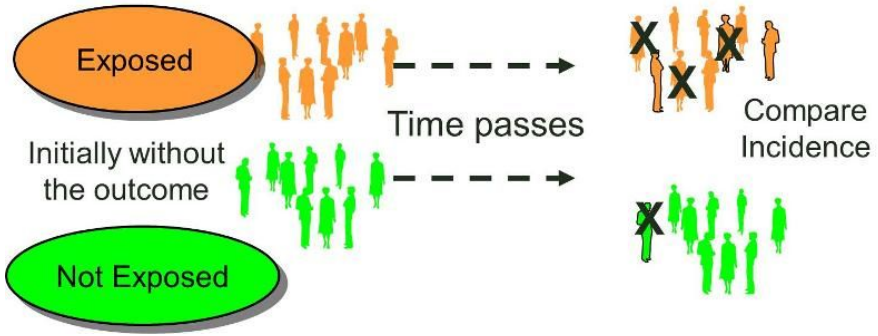
Big data VS Ensaios Clínicos



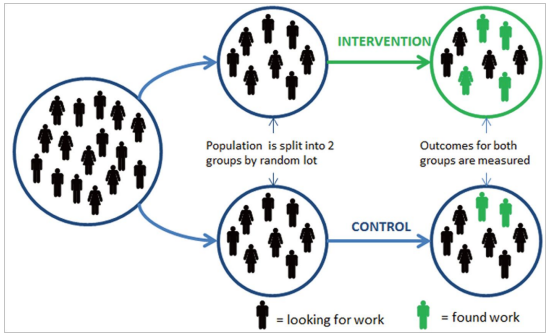
Estudos de associação

Big Data

Coortes, casos e controles



Ensaios



PRECISÃO DOS DADOS, simplicidade da análise, planejamento prévio, erro amostral

Número de associações avaliadas, menor custo, validade externa, velocidade de coleta

Big Data

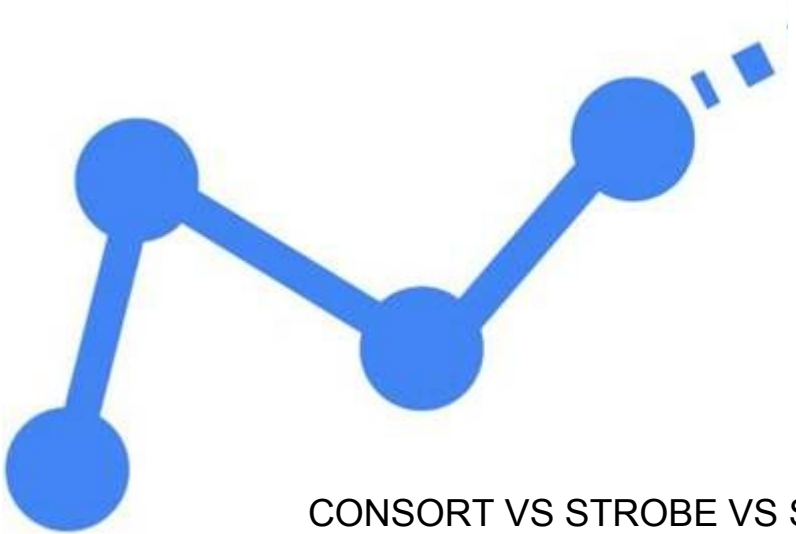
- Associações definitivas*
  - “Ciência vive mudando..”
- Análises hierárquicas
  - evitaria reposição hormonal proteção CV
- Múltiplas descobertas
  - Simultaneas, inesperadas

Big Data

- Mas são observacionais e ..
  - “Correlation is not causation..”
- Dados menos confiáveis
- Maior incidência de associações indiretas
- Muitas vezes os resultados são ininteligíveis..

Para que serve?

Modelos preditores



- Se variável é associada indiretamente ou é efeito não interessa.
- Entender o modelo não interessa.
- Precisão do modelo e consistência (ausência de overfitting) interessam.

CONSORT VS STROBE VS STARD VS TRIPOD

- Cuidado com interpretação de parâmetros de modelos preditores multivariáveis..
- Regressão múltipla é modelo preditor!!!!!!!!!!!!!!!
  - Coeficientes lábeis;
  - Não extrapolar!

Big data VS estatística

- **Frequentista:**
  - **Teorema central do limite… p-valor é ainda menos importante.**
- **Bayesiana**
  - **Pilar na geração/atualização de crenças;**
  - **Interpretação amigável ao usuário.**

IC95 e ICred95

Intervalo 95% = 80 - 100 mmHg (PAM)

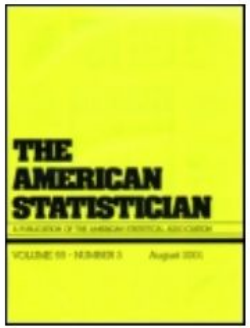
- **Frequentista:**
  - **Se vários estudos forem repetidos, os IC95% da média irão se encontrar nesse intervalo 95% das vezes. Foco = média verdadeira**
- **Bayesiana**
  - **p(PAM 80-100) = 95%. Foco= intervalo esperado**

Estatística em big data

- **Em big data média oscila pouco → artifícios frequentistas são pouco informativos.**
- **Fatores bayesianos (substituem p)**

Fatores bayesianos

- **Interpretação fácil;**



Taylor & Francis
Taylor & Francis Group

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on p-Values: Context, Process, and Purpose

THE AMERICAN STATISTICIAN
2016, VOL. 70, NO. 2, 129–133
http://dx.doi.org/10.1080/00031305.2016.1154108

Fatores bayesianos

- **Interpretação fácil;**

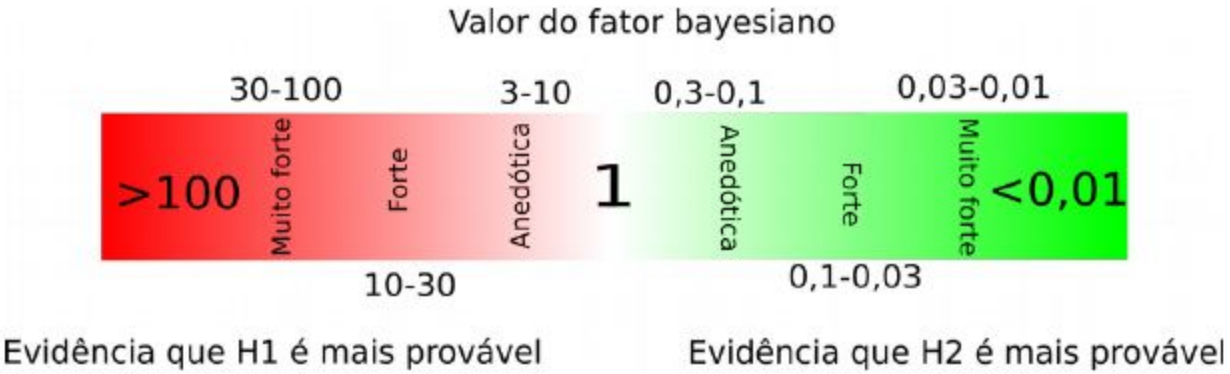


Figura 1: interpretação do valor do fator bayesiano. Neste estudo H1 = probabilidade da complicação ser maior no grupo controle e H2 = probabilidade da complicação ser maior no grupo de estudo.

Fatores bayesianos

- **Interpretação fácil;**
- **Pode comparar 2 modelos diretamente, não precisa tanto de hipótese nula!**

The American Statistician

Taylor & Francis
Taylor & Francis Group

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on p-Values: Context, Process, and Purpose

1. ***P*-values can indicate how incompatible the data are with a specified statistical model.**

THE AMERICAN STATISTICIAN
2016, VOL. 70, NO. 2, 129–133
http://dx.doi.org/10.1080/00031305.2016.1154108

Taylor & Francis
Taylor & Francis Group

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on p-Values: Context, Process, and Purpose

2. ***P*-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.**

The American Statistician

Taylor & Francis
Taylor & Francis Group

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on p-Values: Context, Process, and Purpose

3. **Scientific conclusions and business or policy decisions should not be based only on whether a *p*-value passes a specific threshold.**

that a decision is correct or incorrect. The widespread use of "statistical significance" (generally interpreted as "$p \leq 0.05$") as a license for making a claim of a scientific finding (or implied truth) leads to considerable distortion of the scientific process.

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on p-Values: Context, Process, and Purpose

4. Proper inference requires full reporting and transparency

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on p-Values: Context, Process, and Purpose

5. A *p*-value, or statistical significance, does not measure the size of an effect or the importance of a result.

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on p-Values: Context, Process, and Purpose

6. **By itself, a *p*-value does not provide a good measure of evidence regarding a model or hypothesis.**

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: http://amstat.tandfonline.com/loi/utas20

The ASA's Statement on p-Values: Context, Process, and Purpose

In view of the prevalent misuses of and misconceptions concerning *p*-values, some statisticians prefer to supplement or even replace *p*-values with other approaches. These include methods that emphasize estimation over testing, such as confidence, credibility, or prediction intervals; Bayesian methods; alternative measures of evidence, such as likelihood ratios or Bayes Factors; and other approaches such as decision-theoretic modeling and false discovery rates. All these measures and approaches rely on further assumptions, but they may more directly address the size of an effect (and its associated uncertainty) or whether the hypothesis is correct.

INSTITUTO DO CÂNCER DO ESTADO DE SÃO PAULO
OCTAVIO FARIAS DE OLIVEIRA
ORGANIZAÇÃO SOCIAL DE SAÚDE

EFIC 2013

CBA?

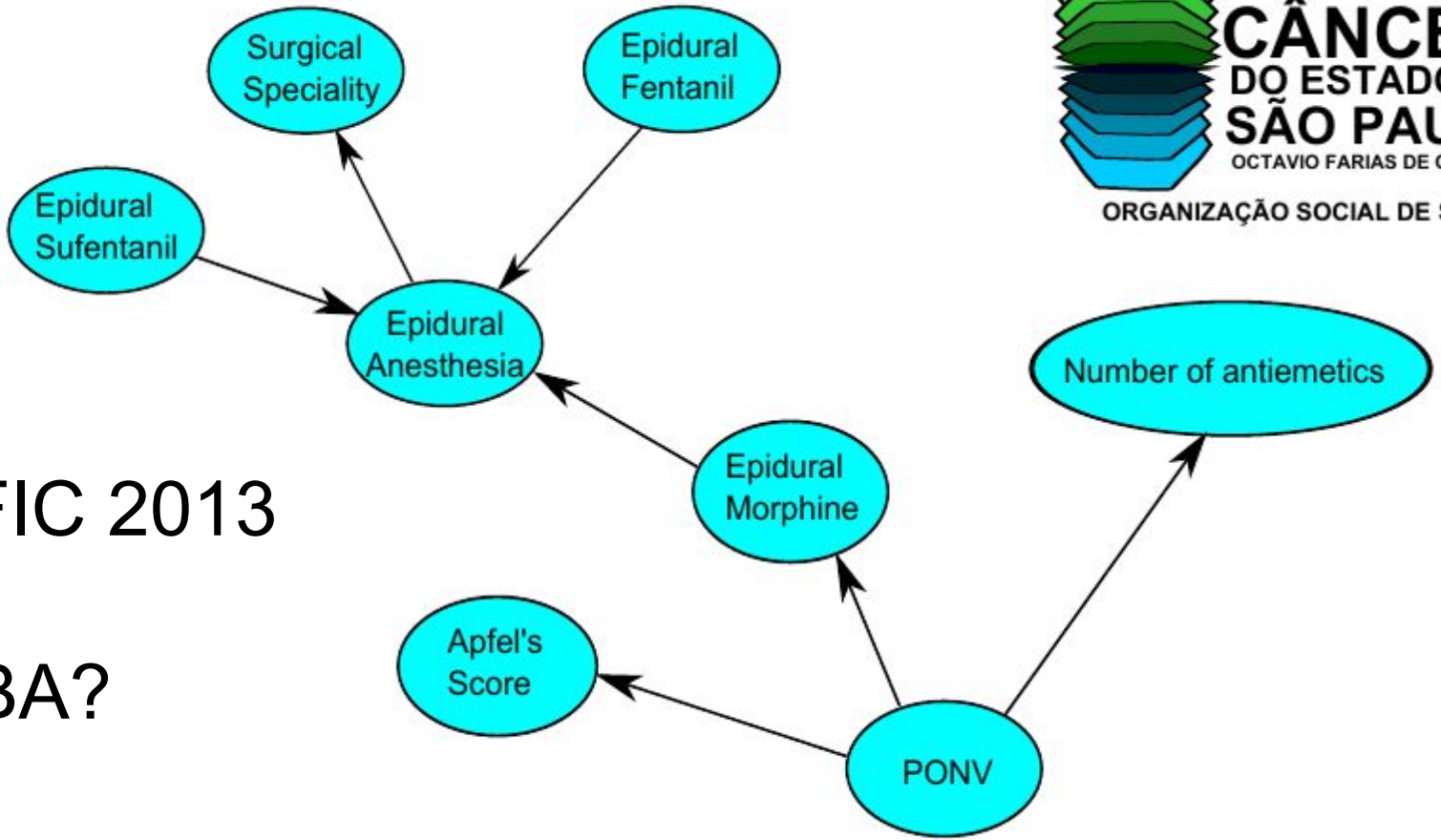
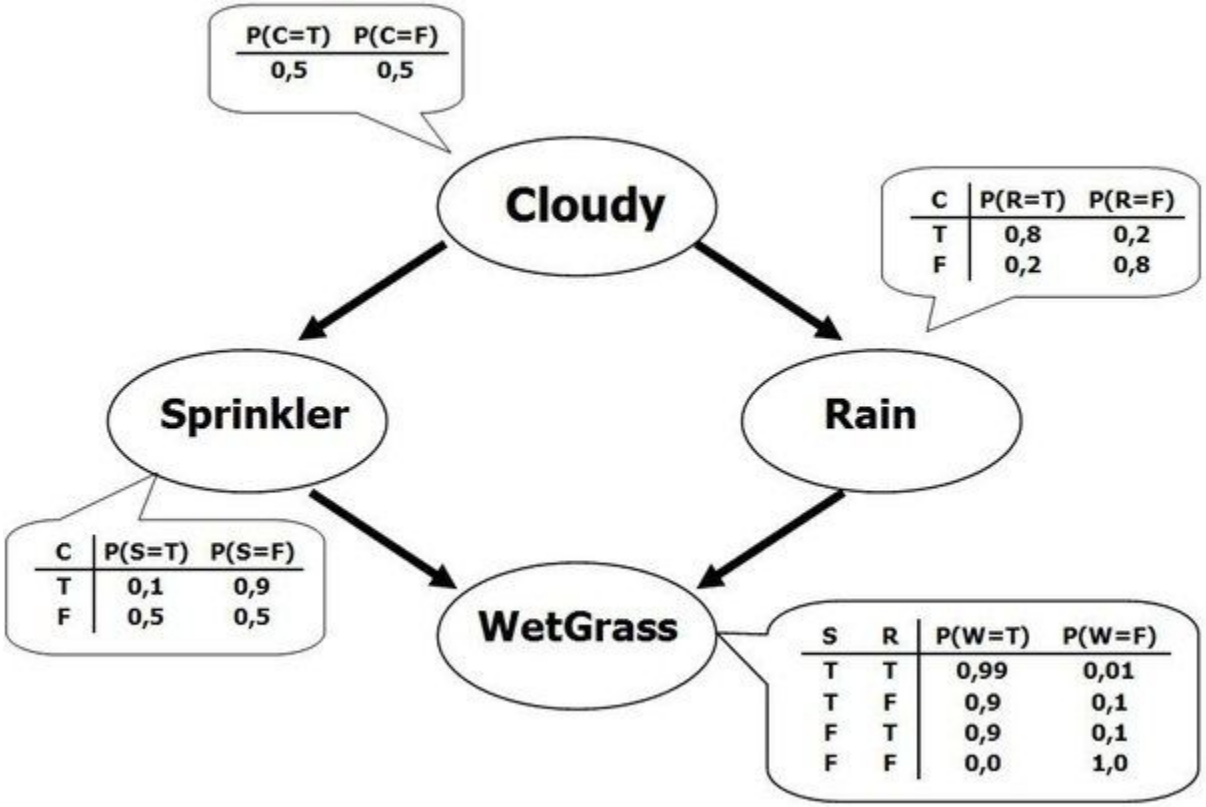


Figure 1: Causal Bayesian Network for the conditional dependence of PONV.


P(C=T) P(C=F)
0,5 0,5
Cloudy
C P(R=T) P(R=F)
T 0,8 0,2
F 0,2 0,8
Sprinkler
Rain
C P(S=T) P(S=F)
T 0,1 0,9
F 0,5 0,5
WetGrass
S R P(W=T) P(W=F)
T T 0,99 0,01
T F 0,9 0,1
F T 0,9 0,1
F F 0,0 1,0

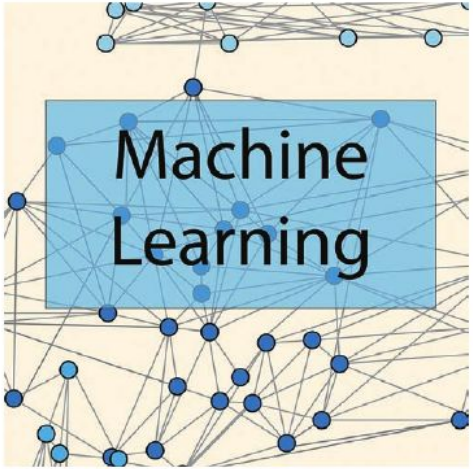
Artificial Intelligence for Everyone

Pedro Gambus, M.D., Steven L. Shafer, M.D.

THIS month's issue of ANESTHESIOLOGY contains a landmark article applying machine learning to model the interaction of remifentanil and propofol on processed electroencephalogram.[1] It's a big deal.

The June 2017 issue of MIT Technology Review claims "Google stakes its future on a piece of software."[2] The software, TensorFlow, was written by Google Brain less than 2 yr ago. TensorFlow is the software behind AlphaGo, a machine learning program that plays the ancient Chinese board game Go. There are approximately 10^{170} Go board positions,[3] vastly exceeding the measly 10^{82} atoms in the observable universe.[4] After a few months of training, AlphaGo readily defeated human Go champions who had trained for a lifetime.

Google is revamping its entire business around TensorFlow and machine learning. In 2016, Google

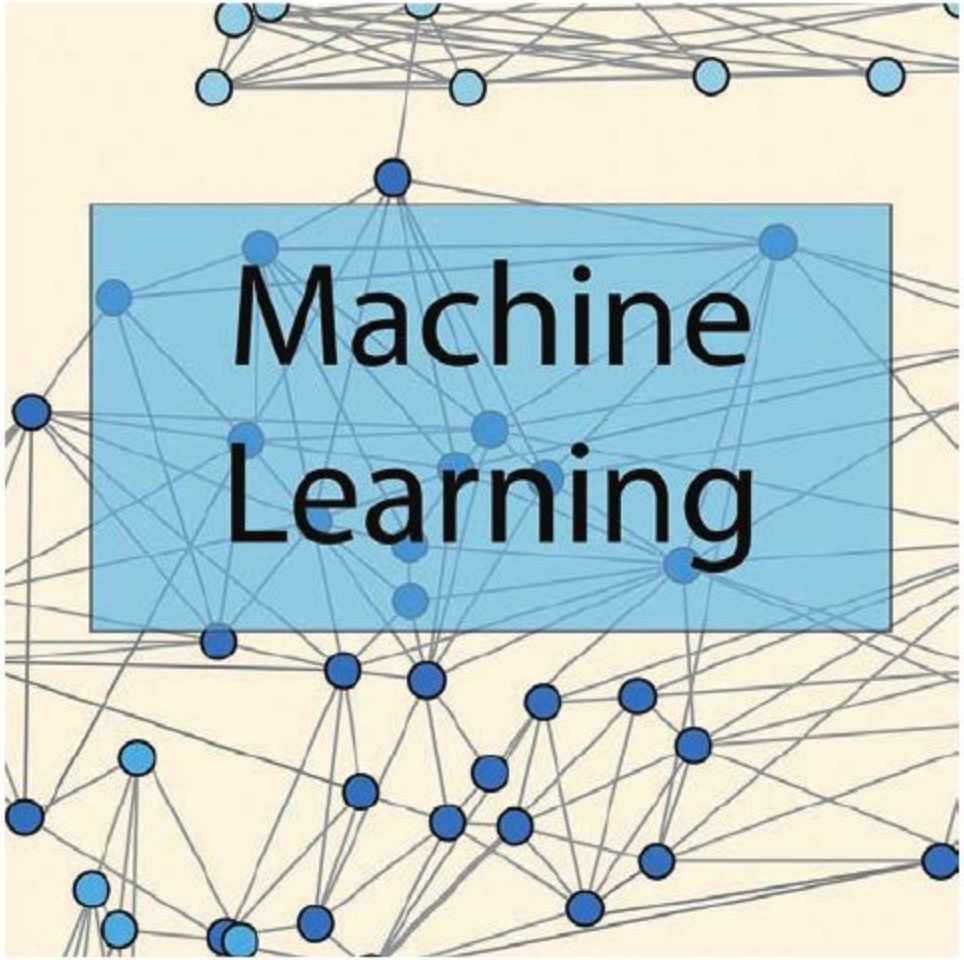


"…machine learning to model the interaction of remifentanil and propofol on processed electroencephalogram…. Is this the end of clinical pharmacology?"

this sphere will become the ruler of the world."[7] Elon Musk quickly agreed: "Competition for AI superiority at national level is the most likely cause of WW3."[8]

What is artificial intelligence and machine learning? Why is Putin praising AI, and Elon Musk fearful? What does AI have to do with neural networks, Tensorflow, and the article by Lee *et al.*?

Artificial intelligence is the capability of a computer to respond in a manner resembling human intelligence.[9] It is nearly impossible to write programs that appear smart. Instead, computers are programed to learn from experience, just like we do. Although some data analysis techniques are described as "learning,"[10] machine learning involves "programming a digital computer to behave in a way which, if done by human beings or animals, would be described as involving the process of learning."[11] As practical matter



Anesthesiology 128(3)
March 2018

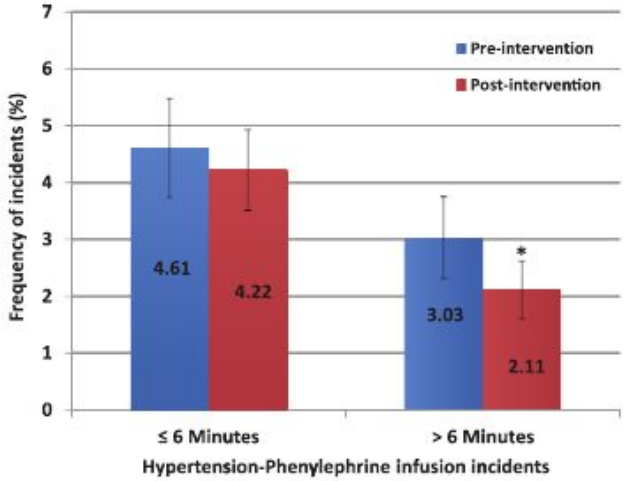
"…machine learning to model the interaction of remifentanil and propofol on processed electroencephalogram…. Is this the end of clinical pharmacology?"

Anesthesia Information Management System-Based Near Real-Time Decision Support to Manage Intraoperative Hypotension and Hypertension

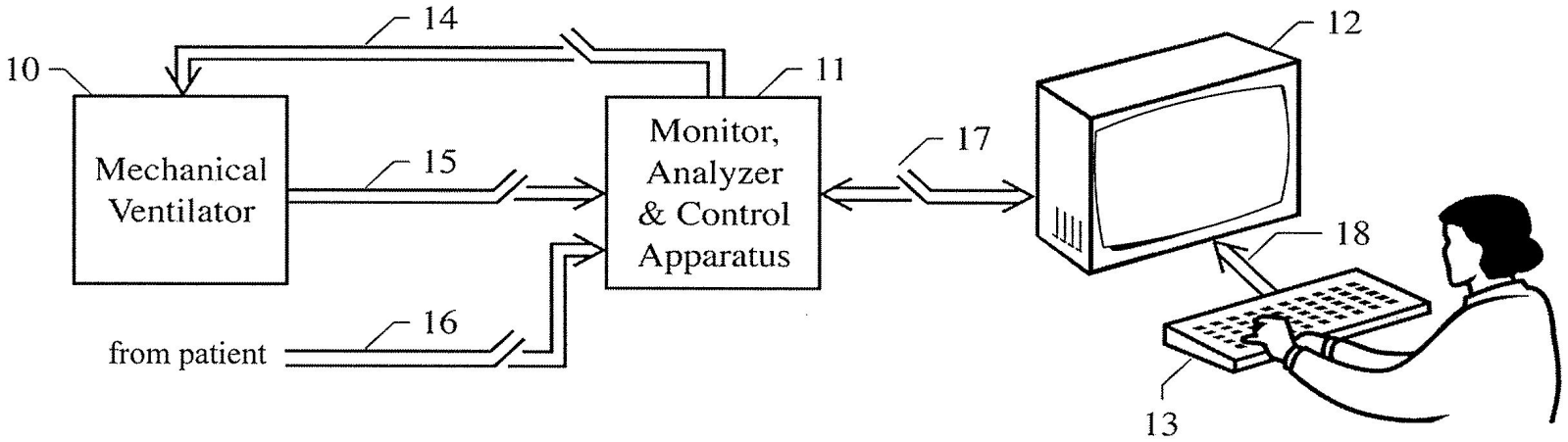
Bala G. Nair, PhD,* Mayumi Horibe, MD,† Shu-Fang Newman, MS,* Wei-Ying Wu, PhD,‡ Gene N. Peterson, MD, PhD,* and Howard A. Schwid, MD*

January 2014 • Volume 118 • Number 1

www.anesthesia-analgesia.org



https://patents.google.com/patent/US8695593B2/en



Weaning and decision support system for mechanical ventilation

The Inevitable Application of Big Data to Health Care

Travis B. Murdoch, MD, MSc
Allan S. Detsky, MD, PhD

gests that 80% of business-related data exist in an unstructured format.[2] The same could probably be said for health care data, a large proportion of which is text-based.

Special Article

Kidney Res Clin Pract 36:3-11, 2017
pISSN: 2211-9132 • eISSN: 2211-9140
https://doi.org/10.23876/j.krcp.2017.36.1.3

KIDNEY RESEARCH AND CLINICAL PRACTICE

Check for updates

Medical big data: promise and challenges

Choong Ho Lee, Hyung-Jin Yoon

Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul, Korea