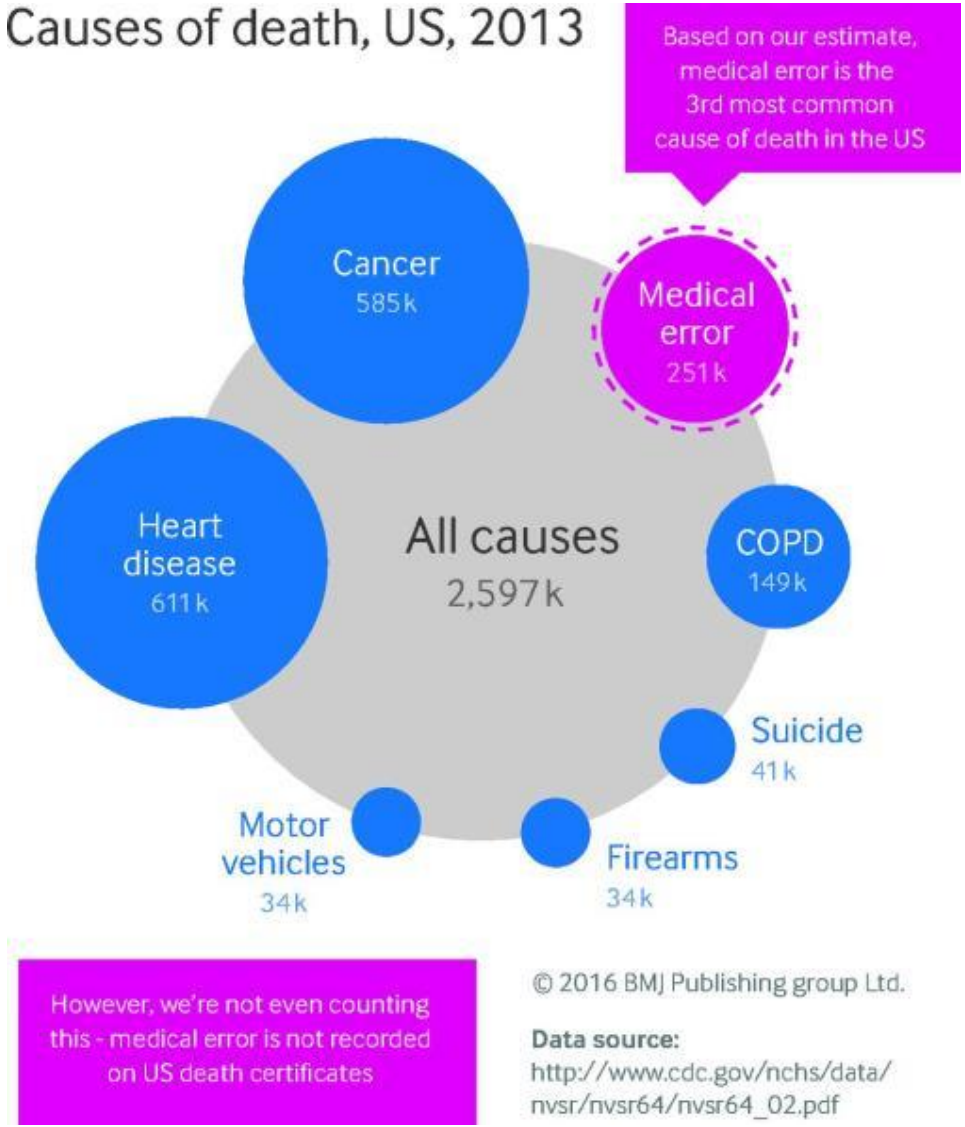
RowAnalytics

transforming the delivery of health

Steve Gardner, CEO

E: steve@rowanalytics.com
T: +44 1865 575170 / +44 7799 671644
W: rowanalytics.com

rowanalytics



thebmj

BMJ 2016;353:i2139 doi: 10.1136/bmj.i2139 (Published 3 May 2016)

ANALYSIS

Medical error—the third leading cause of death in the US

Medical error is not included on death certificates or in rankings of cause of death. **Martin Makary** and **Michael Daniel** assess its contribution to mortality and call for better reporting

Martin A Makary *professor*, Michael Daniel *research fellow*

Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

The annual list of the most common causes of death in the United States, compiled by the Centers for Disease Control and Prevention (CDC), informs public awareness and national research priorities each year. The list is created using death certificates filled out by physicians, funeral directors, medical examiners, and coroners. However, a major limitation of the death certificate is that it relies on assigning an International Classification of Disease (ICD) code to the cause of death.[1] As a result, causes of death not associated with an ICD code, such as human and system factors, are not captured. The science of safety has matured to describe how communication breakdowns, diagnostic errors, poor judgment, and inadequate skill can directly result in patient harm and death. We analyzed the scientific literature on medical error to identify its contribution to US deaths in relation to causes listed by the CDC.[2]

Death from medical care itself

Medical error has been defined as an unintended act (either of omission or commission) or one that does not achieve its intended outcome,[3] the failure of a planned action to be completed as intended (an error of execution), the use of a wrong plan to achieve an aim (an error of planning),[4] or a deviation from the process of care that may or may not cause harm to the patient.[5] Patient harm from medical error can occur at the individual or system level. The taxonomy of errors is expanding to better categorize preventable factors and events.[6] We focus on preventable lethal events to highlight the scale of potential for improvement.

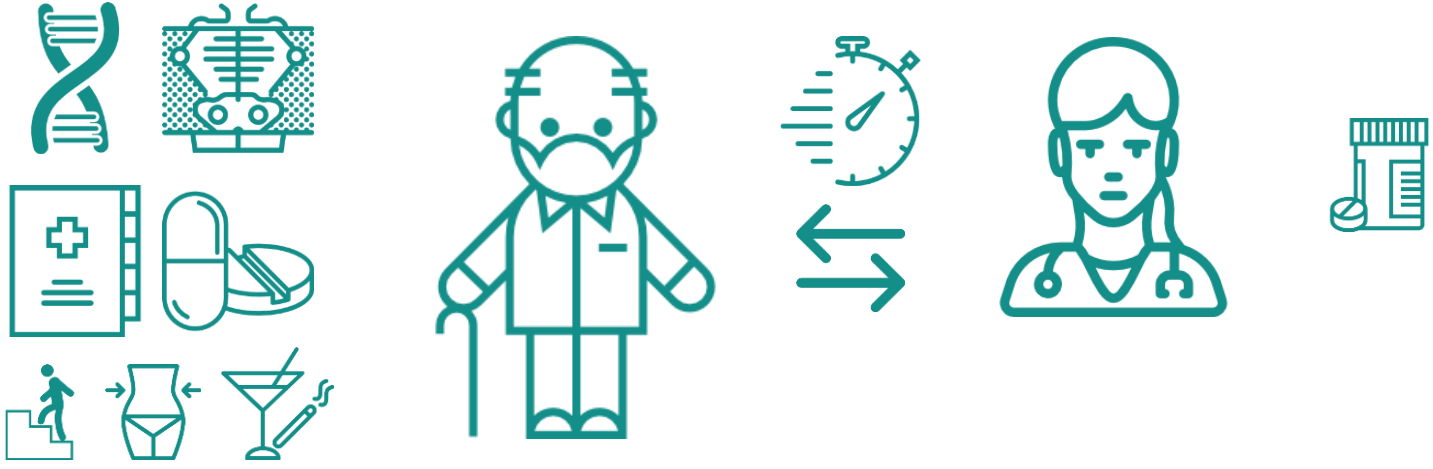
The role of error can be complex. While many errors are non-consequential, an error can end the life of someone with a long life expectancy or accelerate an imminent death. The case in the box shows how error can contribute to death. Moving

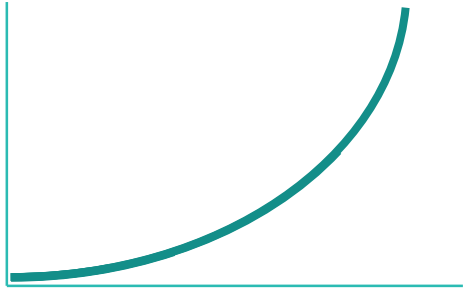
How big is the problem?

The most commonly cited estimate of annual deaths from medical error in the US—a 1999 Institute of Medicine (IOM) report[7]—is limited and outdated. The report describes an incidence of 44 000-98 000 deaths annually.[7] This conclusion was not based on primary research conducted by the institute but on the 1984 Harvard Medical Practice Study and the 1992 Utah and Colorado Study.[8 9] But as early as 1993, Leape, a chief investigator in the 1984 Harvard study, published an article arguing that the study's estimate was too low, contending that 78% rather than 51% of the 180 000 iatrogenic deaths were preventable (some argue that all iatrogenic deaths are preventable).[10] This higher incidence (about 140 400 deaths due to error) has been supported by subsequent studies which suggest that the 1999 IOM report underestimates the magnitude of the problem. A 2004 report of inpatient deaths associated with the Agency for Healthcare Quality and Research Patient Safety Indicators in the Medicare population estimated that 575 000 deaths were caused by medical error between 2000 and 2002, which is about 195 000 deaths a year (table 1⇓).[11] Similarly, the US Department of Health and Human Services Office of the Inspector General examining the health records of hospital inpatients in 2008, reported 180 000 deaths due to medical error a year among Medicare beneficiaries alone.[12] Using similar methods, Classen et al described a rate of 1.13%.[13] If this rate is applied to all registered US hospital admissions in 2013[15] it translates to over 400 000 deaths a year, more than four times the IOM estimate.

Similarly, Landrigan et al reported that 0.6% of hospital admissions in a group of North Carolina hospitals over six years (2002-07) resulted in lethal adverse events and conservatively

The Personalization of Medicine



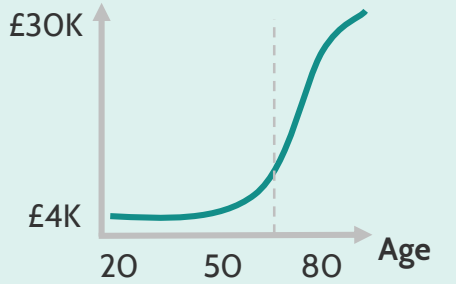


precision.life

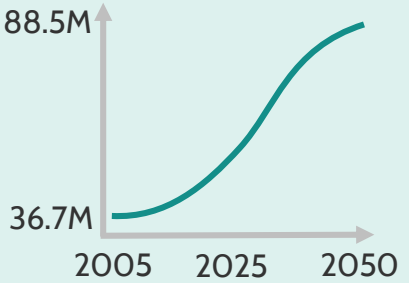
What can Precision Medicine Deliver?

Patients with long-term conditions on 3+ drugs cost 5x more and have poorer outcomes

US Health costs/yr per person[1]



US population aged 65+ [2]



Precision medicine can deliver:

48% fewer ER visits[3,6]

40% fewer readmissions[3,6]

75% better patient reported outcomes[4]

£225B+/yr compliance related savings to US healthcare[5,6,7]

[1] CMS.gov (2014)
[2] Kaiser Family Foundation (2015) The Rising Cost of Living Longer
[3] Clinical impact of pharmacogenetic profiling (2017) Elliott LS et al. PLoS ONE 12(2)
[4] Walsh & Cussen, Ir Med J. 2010 103(8):236-8
[5] Viswanathan et al Interventions to Improve Adherence (2012) Annals of Internal Med
[6] Sultana, J., Cutroneo, P., Trifioro, G. (2013). J. Pharmacology & Pharmacotherapeutics
[7] Brixner, D., et al Effect of pharmacogenetic profiling (2015). J Med Economics 19, 3

precision.life

Healthcare Industry Megatrends

Outcomes	• Payment on results – more accurate diagnosis & evidence of cost-effective patient benefit
Precision Medicine	• Giving the right treatments to the right patient at the right time, every time
Patient Engagement	• Helping patients & carers be better informed and actively change behavior to manage their health

precision.life

Whole Genome Analysis

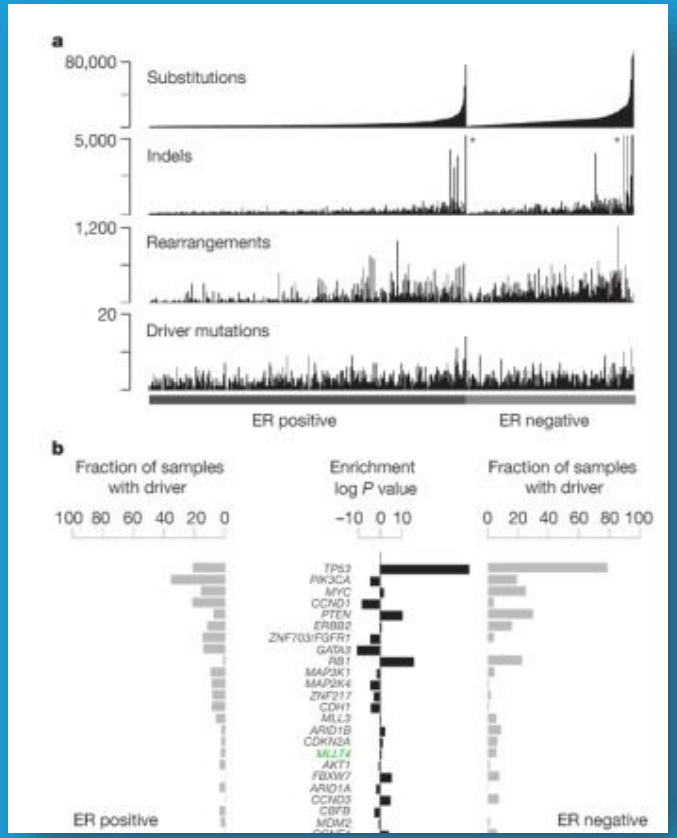
ARTICLE

doi:10.1038/nature17676

Landscape of somatic mutations in 560 breast cancer whole-genome sequences

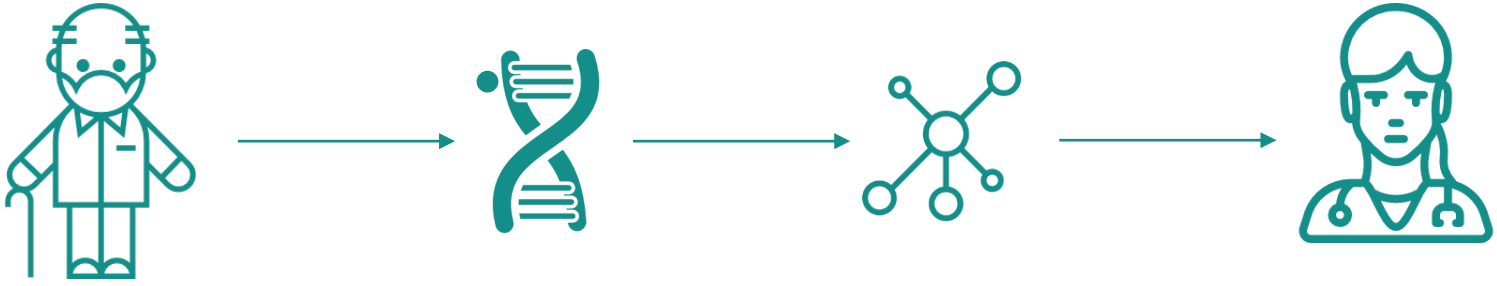
Serena Nik-Zainal[1,2], Helen Davies[1], Johan Staaf[3], Manasa Ramakrishna[1], Dominik Glodzik[1], Xueqing Zou[1], Inigo Martincorena[1], Ludmil B. Alexandrov[1,4,5], Sancha Martin[1], David C. Wedge[1], Peter Van Loo[1,6], Young Seok Ju[1], Marcel Smid[7], Arie B. Brinkman[8], Sandro Morganella[9], Miriam R. Aure[10,11], Ole Christian Lingjærde[11,12], Anita Langerød[10,11], Markus Ringnér[3], Sung-Min Ahn[13], Sandrine Boyault[14], Jane E. Brock[15], Annegien Broeks[16], Adam Butler[1], Christine Desmedt[17], Luc Dirix[18], Serge Dronov[1], Aquila Fatima[19], John A. Foekens[7], Moritz Gerstung[1], Gerrit K. J. Hooijer[20], Se Jin Jang[21], David R. Jones[1], Hyung-Yong Kim[22], Tari A. King[23], Savitri Krishnamurthy[24], Hee Jin Lee[25], Jeong-Yeon Lee[25], Yilong Li[1], Stuart McLaren[1], Andrew Menzies[1], Ville Mustonen[1], Sarah O'Meara[1], Iris Pauporté[26], Xavier Pivot[27], Colin A. Purdie[28], Keiran Raine[1], Kamna Ramakrishnan[1], F. Germán Rodríguez-González[7], Gilles Romieu[29], Anieta M. Sieuwerts[7], Peter T. Simpson[30], Rebecca Shepherd[1], Lucy Stebbings[1], Olafur A. Stefansson[31], Jon Teague[1], Stefania Tommasi[32], Isabelle Treilleux[33], Gert G. Van den Eynden[18,34], Peter Vermeulen[18,34], Anne Vincent-Salomon[35], Lucy Yates[1], Carlos Caldas[36], Laura van't Veer[16], Andrew Tutt[37,38], Stian Knappskog[39,40], Benita Kiat Tee Tan[41,42], Jos Jonkers[16], Åke Borg[3], Naoto T. Ueno[24], Christos Sotiriou[17], Alain Viari[43,44], P. Andrew Futreal[1,45], Peter J. Campbell[1], Paul N. Span[46], Steven Van Laere[18], Sunil R. Lakhani[30,47], Jorunn E. Eyfjord[31], Alastair M. Thompson[28,48], Ewan Birney[9], Hendrik G. Stunnenberg[8], Marc J. van de Vijver[20], John W. M. Martens[7], Anne-Lise Børresen-Dale[10,11], Andrea L. Richardson[15,19], Gu Kong[22], Gilles Thomas[44] & Michael R. Stratton[1]

We analysed whole-genome sequences of 560 breast cancers to advance understanding of the driver mutations conferring clonal advantage and the mutational processes generating somatic mutations. We found that 93 protein-coding cancer genes carried probable driver mutations. Some non-coding regions exhibited high mutation frequencies, but most have distinctive structural features probably causing elevated mutation rates and do not contain driver mutations. Mutational signature analysis was extended to genome rearrangements and revealed twelve base substitution and six rearrangement signatures. Three rearrangement signatures, characterized by tandem duplications or deletions, appear associated with defective homologous-recombination-based DNA repair: one with deficient BRCA1 function, another with deficient BRCA1 or BRCA2 function, the cause of the third is unknown. This analysis of all classes of somatic mutation across exons, introns and intergenic regions highlights the repertoire of cancer genes and mutational processes operating, and progresses towards a comprehensive account of the somatic genetic basis of breast cancer.



precision.life

Precision Medicine Use Case



SEQUENCE
patient
genome

IDENTIFY
mutations
(SNPs)

ANALYZE
metabolic
function

DECIDE
best drug
prescription

precision.life

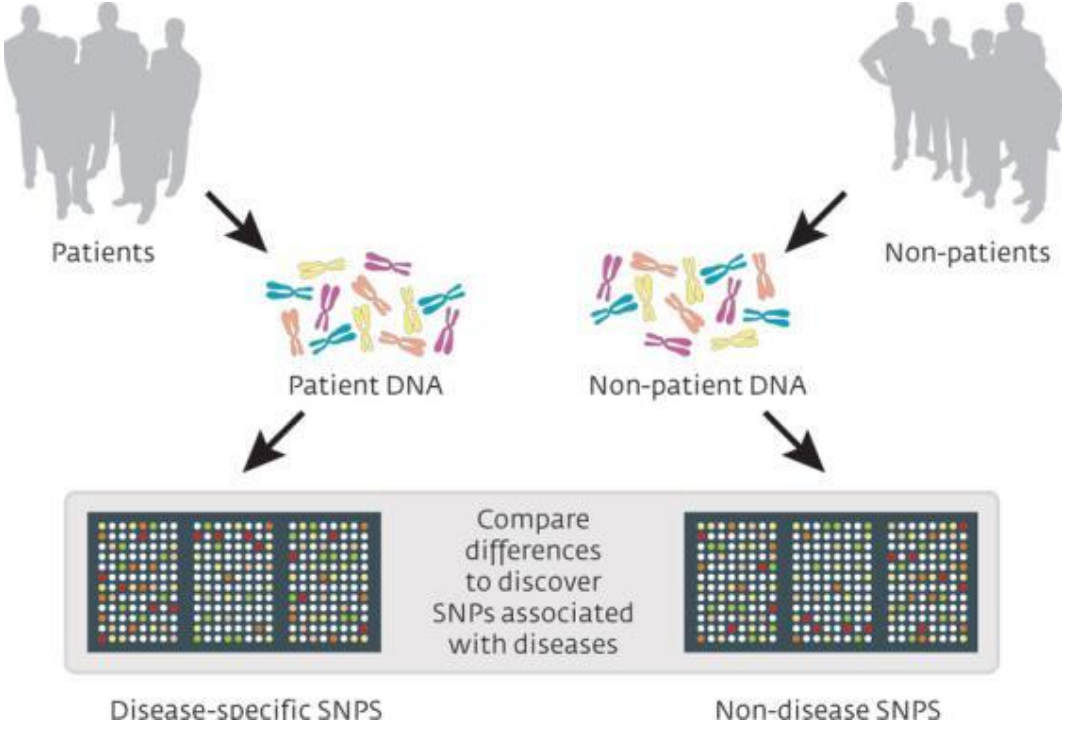
World's fastest and most scalable genome association studies

Powerful and efficient deep semantic learning & search tools

Most personalized clinical decision support & digital health platform

‘Traditional’ Biomarker Discovery / GWAS

- Correlate genetic markers with disease/treatment outcomes



precision.life

Source: MMG 233 2014 Genetics & Genomics Wiki

The Biological Challenge

Leading Edge
Perspective

Cell

An Expanded View of Complex Traits: From Polygenic to Omnigenic

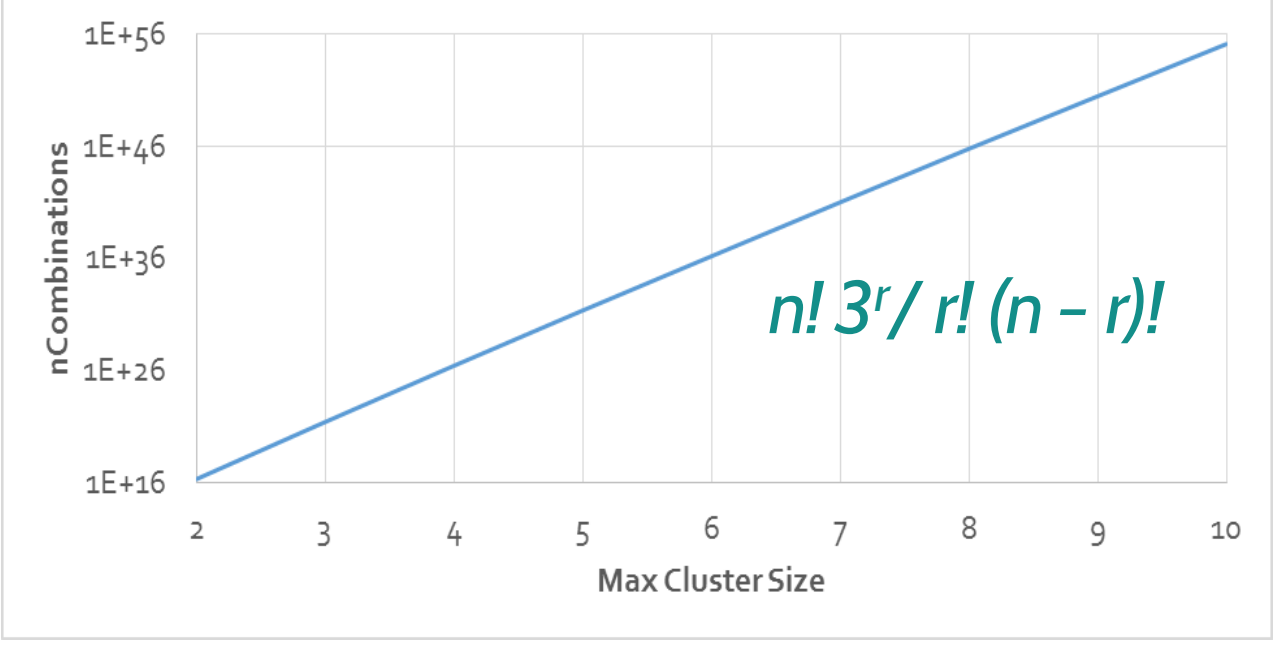
Evan A. Boyle,[1,*] Yang I. Li,[1,*] and Jonathan K. Pritchard[1,2,3,*]
[1]Department of Genetics
[2]Department of Biology
[3]Howard Hughes Medical Institute
Stanford University, Stanford, CA 94305, USA
*Correspondence: eaboyle@stanford.edu (E.A.B.), yangili@stanford.edu (Y.I.L.), pritch@stanford.edu (J.K.P.)
http://dx.doi.org/10.1016/j.cell.2017.05.038

A central goal of genetics is to understand the links between genetic variation and disease. Intuitively, one might expect disease-causing variants to cluster into key pathways that drive disease etiology. But for complex traits, association signals tend to be spread across most of the genome—including near many genes without an obvious connection to disease. We propose that gene regulatory networks are sufficiently interconnected such that all genes expressed in disease-relevant cells are liable to affect the functions of core disease-related genes and that most heritability can be explained by effects on genes outside core pathways. We refer to this hypothesis as an "omnigenic" model.

rowanalytics

The Computational Challenge

- Current fastest supercomputer does 3 x 10^{16} operations/sec
  - At n = 6 would take 3.1 trillion years...
  - At n = 10 would take 3.6 x 10^{38} seconds



precision.life

Bipolar Study Findings

Layer (# SNPs in combination)	# Networks	Features	% Cases (cumulative)	% Controls (cumulative)
1	-		-	-
2	-		-	-
3	3	Rare variant homozygote		0
4	1	Rare variant homozygote	26% (158/607)	0
5				-
6				-
7				-
8				-
9				-
10	1	Common variant homozygote/ heterozygote	44% (222/607)	0

607
Bipolar patients

1,355
Controls

803
SNPs per person

1.7×10^{28}
possible combinations

The n-SNP networks are genetically distinct and have been clinically validated

precision.life

Genomics Data Representation

SNP Genotype

247^0

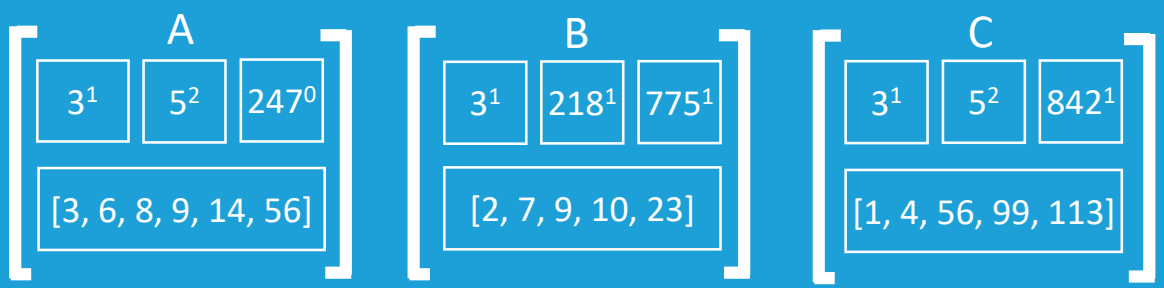
SNP Genotype:
0 = homozygous 'normal/major allele'
1 = heterozygous
2 = homozygous 'variant/minor allele'

SNP Index 247 = rs12345678

Case #27 Genotype

1^0 2^0 3^1 4^0 5^2 6^1 ... n^0

States
Layer 3



(SNP Genotypes)

(Case Indices)

where nCases => minCases (e.g. 5 above)
and nControls <= maxControls (e.g. 0)

precision.life

Synomics Example – Breast Cancer Study

14,777
People with
BRCA1/2 mutations

200K
SNPs per person

- All participants have BRCA 1 and/or BRCA 2 mutations
  - 3,850 affected by breast cancer (cases)
  - 10,927 non-affected (controls)
- Seeking combinations of multiple SNPs associated with:
  - disease risk
  - disease protective effect
  - therapy response

precision.life

Synomics – Transforming Genomic Medicine

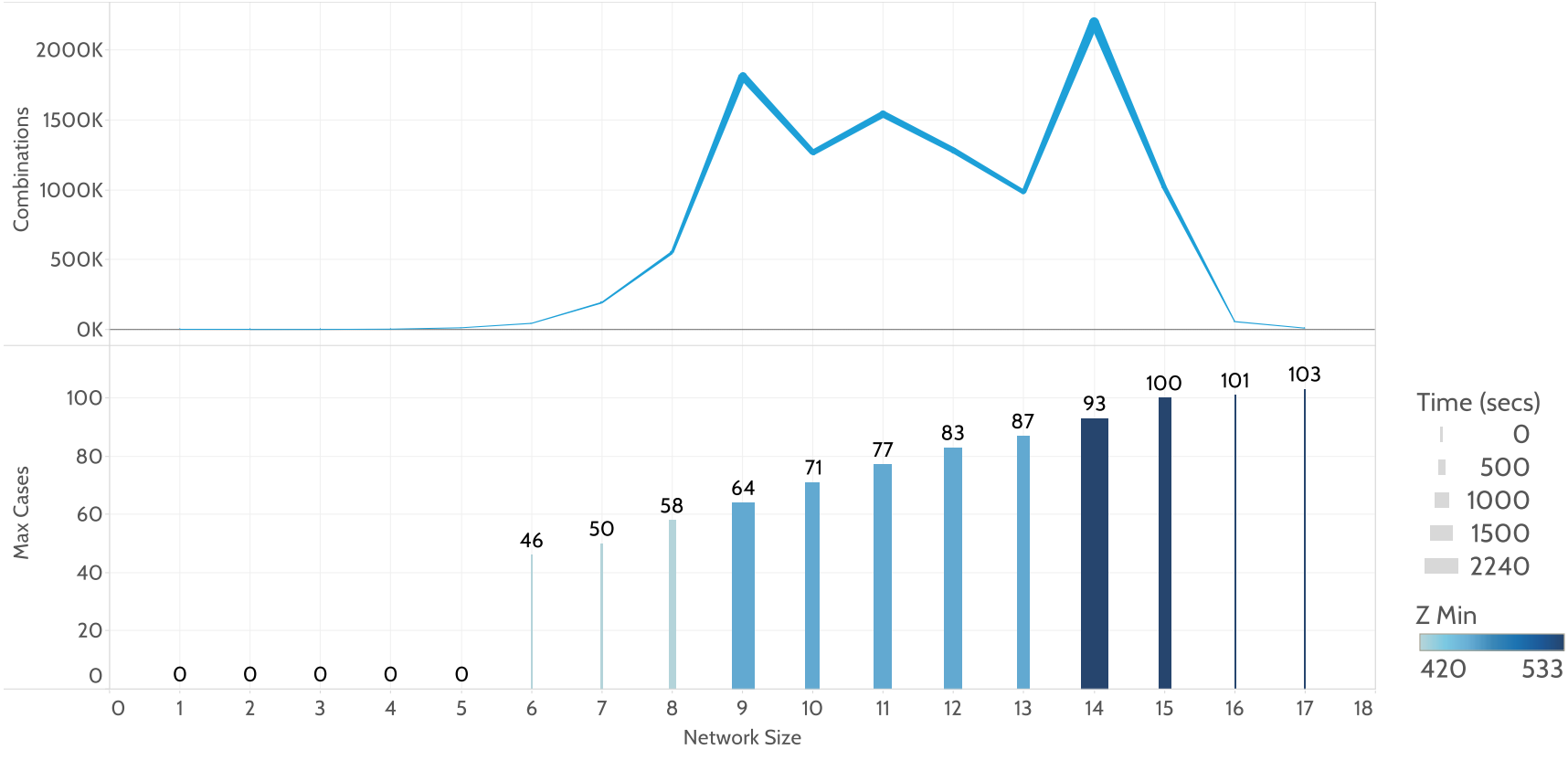
14,777
People with BRCA1/2 mutations

200K
SNPs per person

	Comb.	Current GWAS (1,000 node supercomputer)	**synomics (on single IBM Minsky with 4x Nvidia P100)**
2 SNPs	10^{11}	6-8 months	**12 mins**
6 SNPs	10^{32}	-	**6 hours**
17 SNPs	10^{84}	-	**6 days**

precision.life

Synomics - Breast Cancer (BRCA1)



precision.life

BRCA2 affected / non-affected

1,576
Affected (cases)

6,402
Non-Affected (controls)

200K
SNPs per person

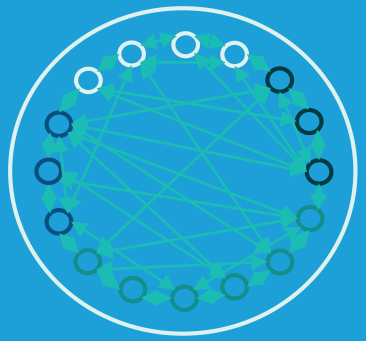
False Discovery Rate (%)*	SNP Genotypes	Cases	Penetrance (%)
20	2,113	799	50.7
10	1,320	627	39.8
5	868	513	32.6
1	142	221	14.0

* Using Benjamini-Hochberg correction for multiple testing

precision.life

Synomics Example – Breast Cancer Study

Key Findings:



Found co-occurring sub-clusters of 3, 4, 5, or 6 SNP variants

- co-occurring in later layers of analysis (8 SNPs+)
- SNPs associated with same pathways show disease functional units
- opportunity to identify combinatorial therapies

Detected 17 SNP networks in up to 103 cases and 0 controls
Very high (>25%) penetrance for good clinical relevance

Identified disease protective & disease risk associated factors

- BRCA1/2 status may suggest risk, but other variants in combination confer an overall greater protective effect

Currently analysing phenotypic and clinical features

precision.life

Genomics Data Representation

Merged Networks

(based on sharing at least x% SNP genotypes)

3^1	[A, B, C, D]	218^1	[B, E]
5^2	[A, C]	842^1	[C, F]
247^0	[A, D]	775^1	[B, F]
264^1	[D, E]	511^2	[E, F]

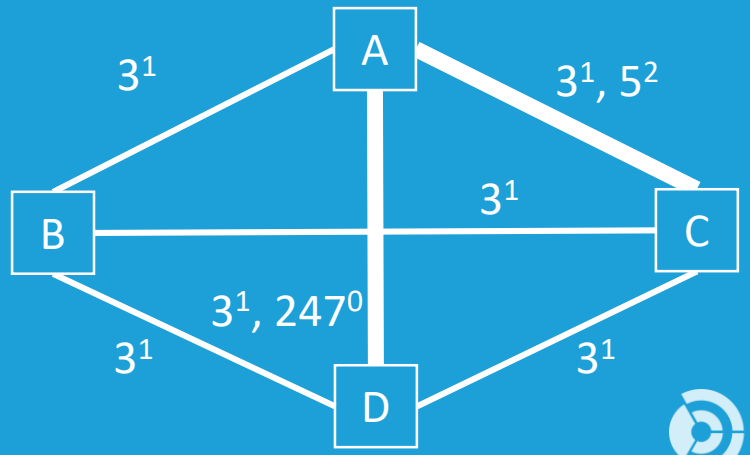
3^1	[A, B, C, D]
5^2	[A, C]
247^0	[A, D]

merged cases

[1, 2, 3, 4, 6, 7, 8, 9, 10, 14, 23, 56, 99, 113, 246, 299, ...]

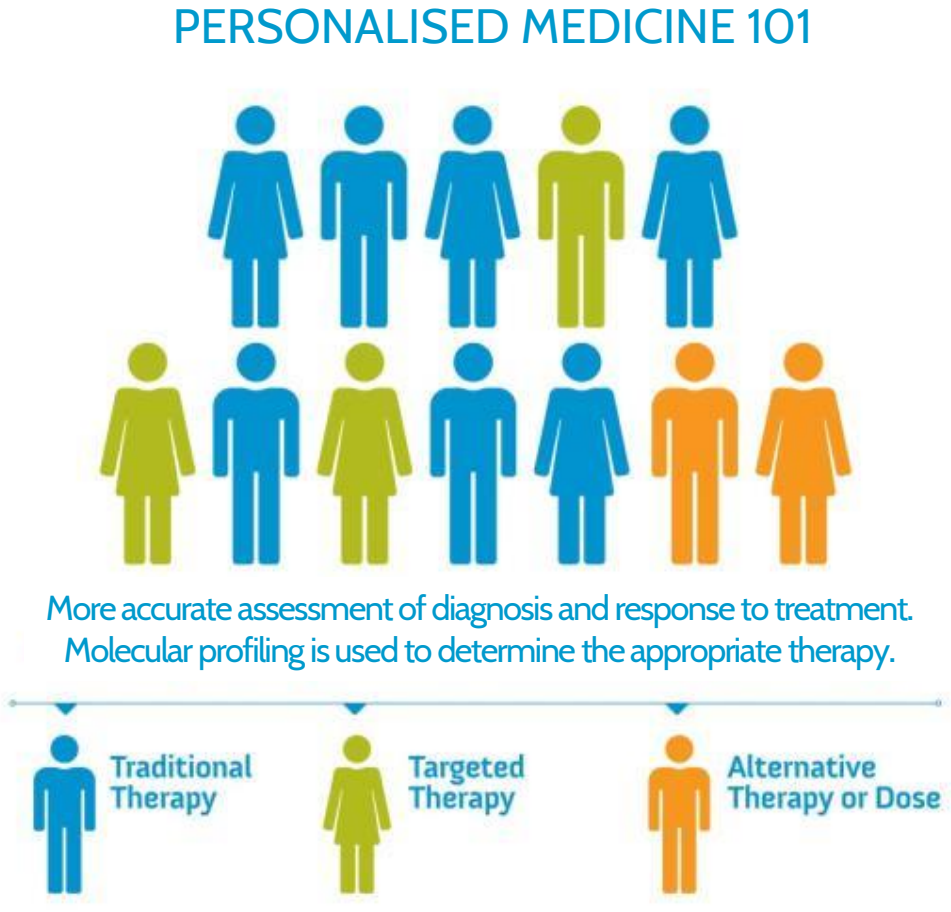
Merged Networks

higher nCases with lower nSGs and high densities is better, i.e. a small number of highly interconnected SNP genotypes



precision.life

Real World Personalisation Challenges



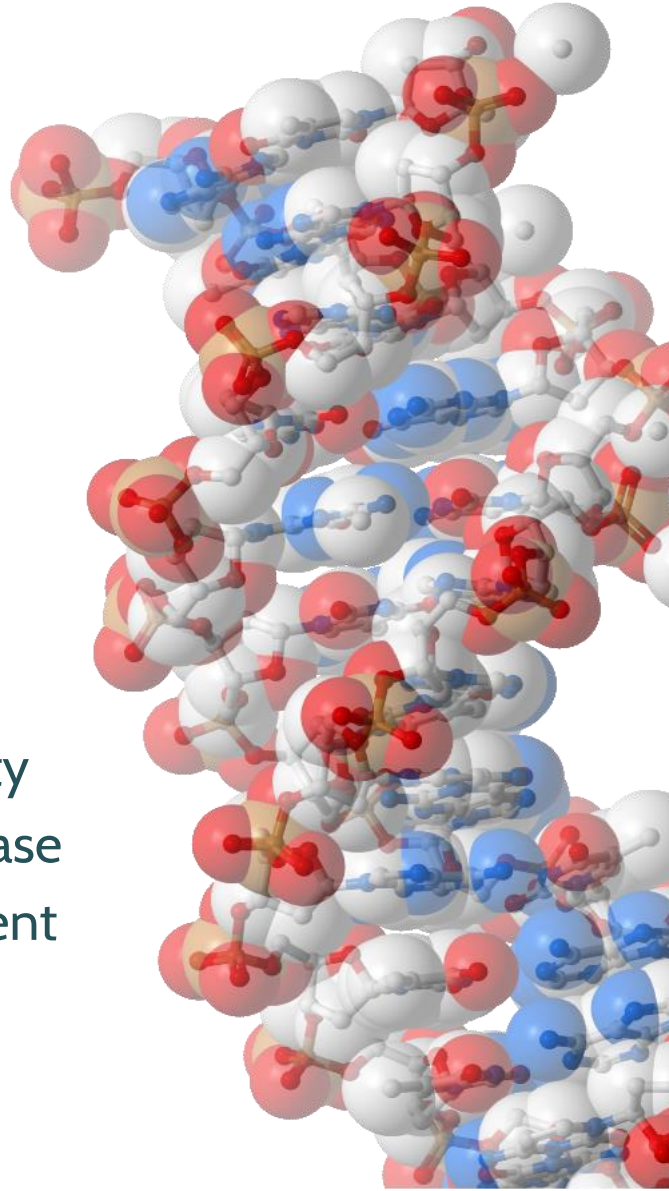
Adapted from: sarahcannon.com

- Clinical status
- Phenotype
- Co-morbidities
- Co-prescriptions
- Lifestyle & environment
- (Clonal) heterogeneity
- Polygenic disease aetiology

precision.life

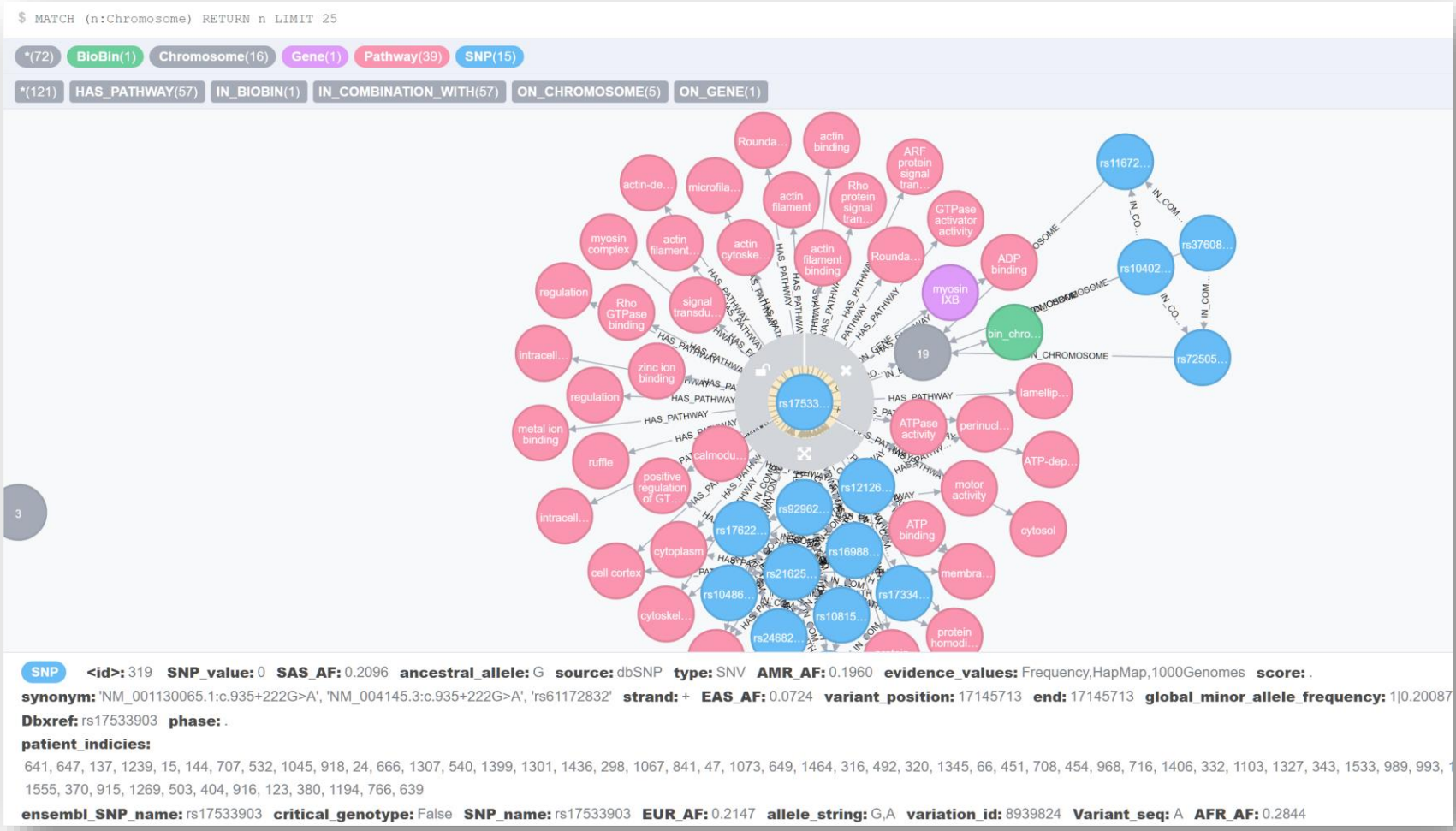
project MinE
Make it yours

- 15,000 MND patients / 7,500 controls
  - 40% whole genomes sequenced
  - 2 petabytes
- Multi-factor late-onset disease
  - Only 5-10% genetically determined heritability
  - 6 independent factors required to trigger disease
  - Imaging, epigenetics, lifestyle, diet, environment clinical history, co-morbidities



precision.life

Biological Interpretation



Biological Annotation

spot.my BETA
LIFE SCIENCE

- **Query:** rs3734805 rs9383935 rs9383589 c6_pos151989450 rs4648881 rs9383936 CD14+CD16- monocyte CD8+/ab T fetal thymus naive B cell
  - Full context of all cell types in which epigenetic activation occurs
  - Literature search (keywords) gave no relevant results (Google/PubMed)
  - Deep semantic search identified 36 relevant papers including:
    "In vivo modulation of the distribution of thymocyte subsets: effects of estrogen..."

- Further queries identified a study where female infants with enlarged thymus treated with X-rays were observed to have higher incidence of breast cancer 36 years later
- Suggested novel disease sensitization mechanism

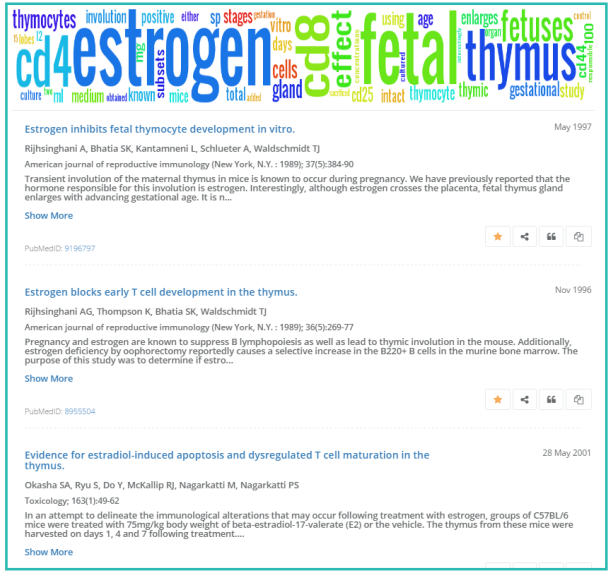
spot.my – deep semantic matching

use keywords or whole paragraphs to search

- matches all of the words to find the best, most relevant hits
- more words = better context = better hits

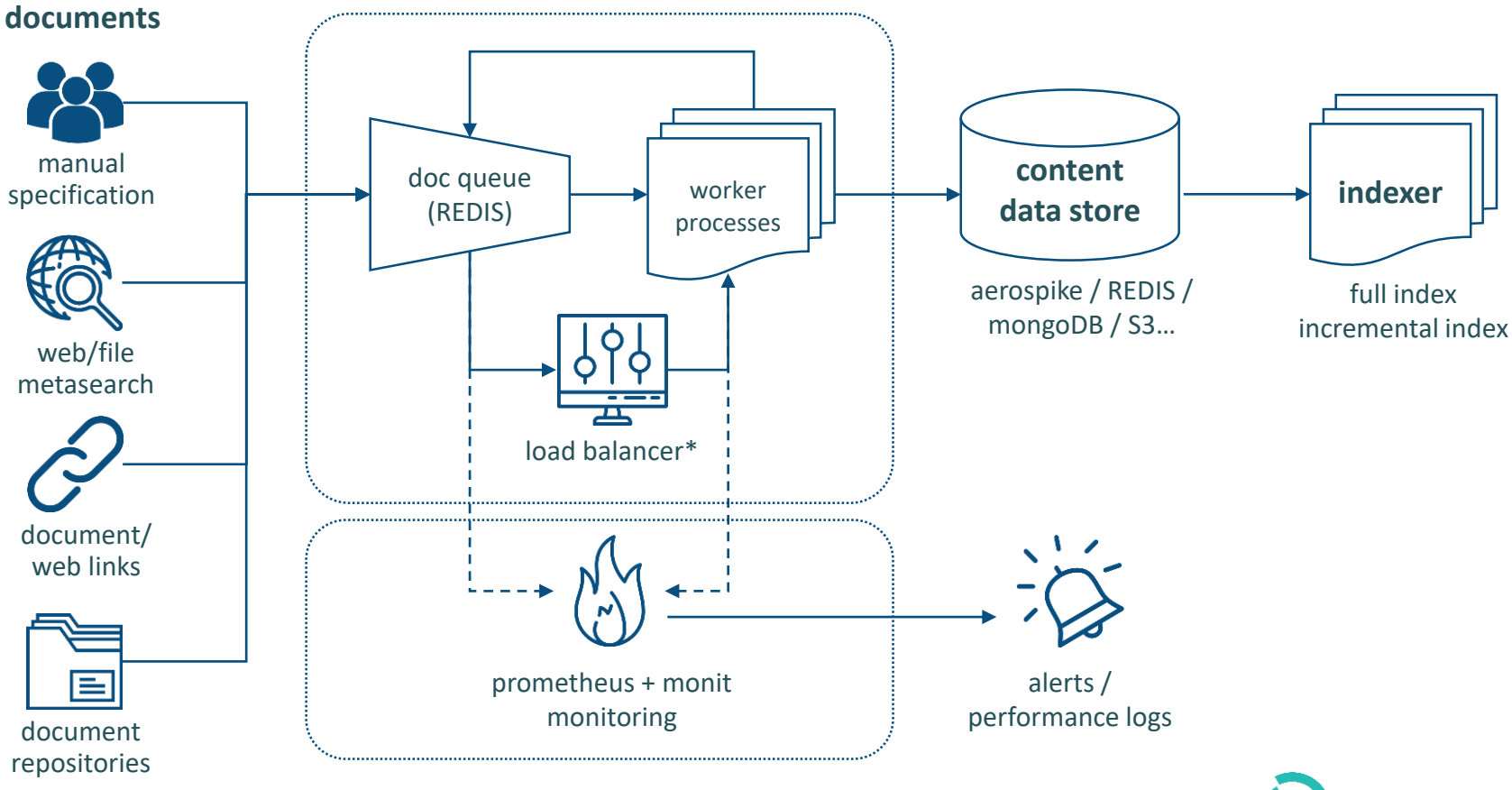
"Transient involution of the maternal thymus in mice is known to occur during pregnancy. Although estrogen crosses the placenta, fetal thymus gland enlarges with advancing gestational age. It is not known if fetal thymocytes are resistant to estrogen or if there are other factors that prevent estrogen from exerting an effect on the development of fetal thymocytes. Therefore we studied the effect of estrogen on isolated fetal thymic glands in vitro. All CD4 and CD8 defined T cell subsets were reduced with a disproportionate loss of CD4+ single positive (SP), CD8+ SP: CD4+CD8+ double positive (DP) cells."

Search



spot.my BETA
LIFE SCIENCE

spot.my GPU enabled semantic search

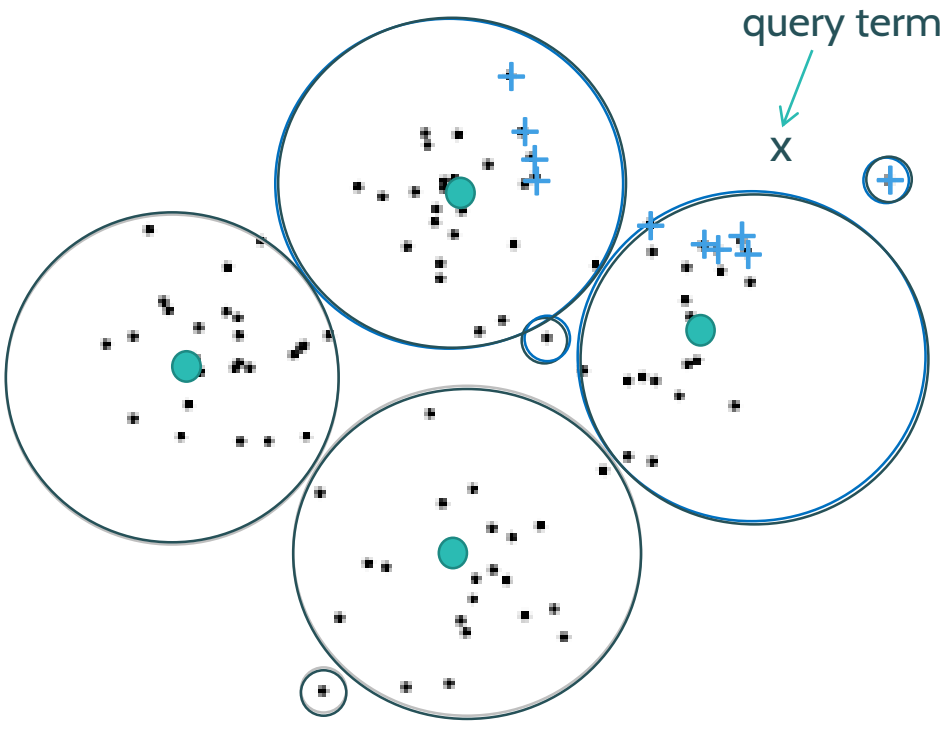


* scraping & indexing infrastructure is fully scalable and distributed
spot.my automatically scales to meet demand and is fully monitored with failure alerts

precision.life

spot.my GPU enabled semantic search

- Enables very fast searching of large corpora & vocabs with low RAM/CPU



STAGE 1 (cluster articles)

Identify clusters of similar stories and orphans (non-clustered)

27M papers > 5M clusters + 3M orphans
GPU with fast GRAM

STAGE 2 (cluster search)

Find closest clusters and orphans
GPU with fast GRAM

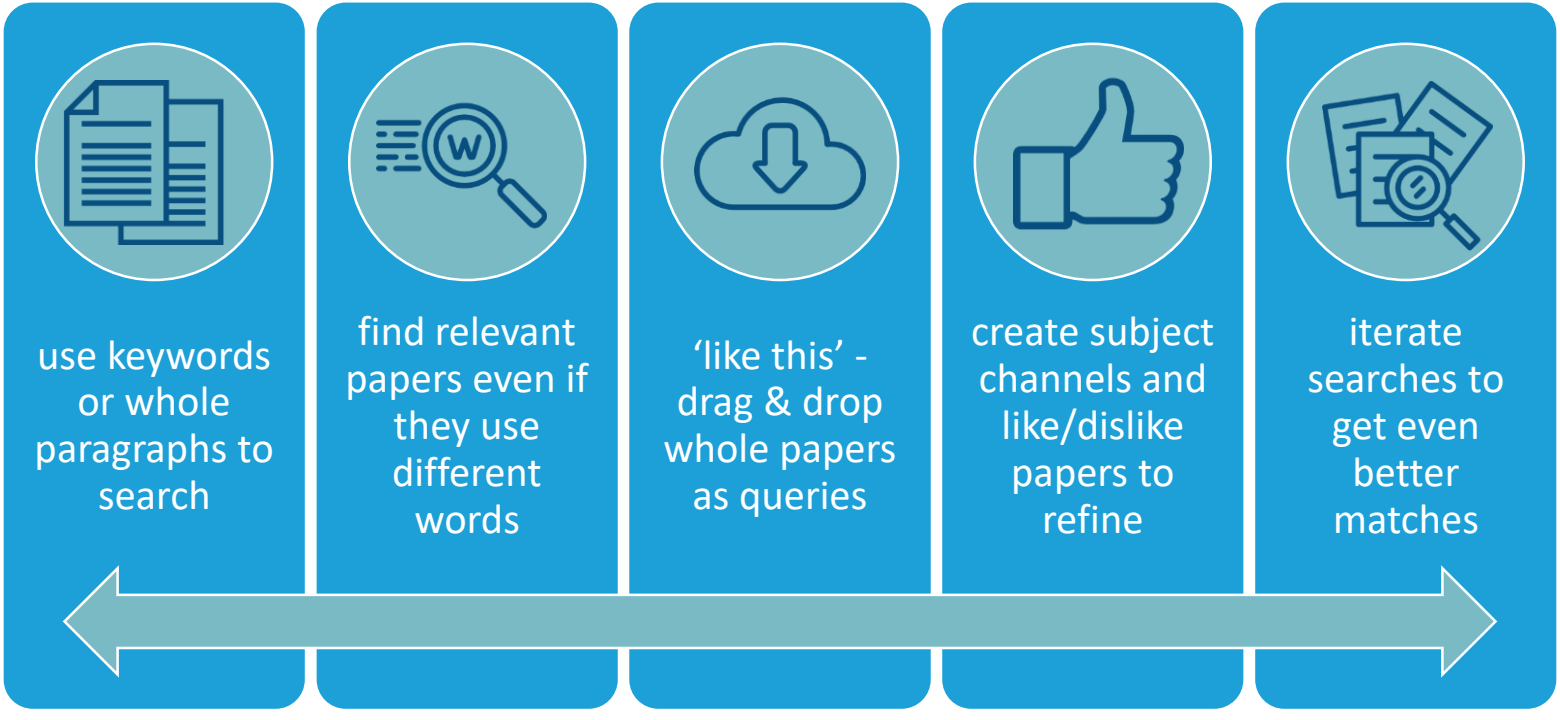
STAGE 3 (results refinement)

Full search of selected clusters to find hits
CPU & cheap RAM

precision.life

other spot.my features

spot.my BETA
LIFE SCIENCE



Creating New Opportunities

Much deeper insight into complex diseases

Novel (patentable) R&D / combinatorial interventions

Includes genotypic, phenotypic and clinical data

Clinical trials design / patient stratification

Healthcare analytics / service planning

Use of biomarker clusters in clinical decision support:

Personalized disease risk scoring and therapy selection

Personalized dietary and lifestyle advice

precision.life

Example - Meet Albert

- High cholesterol, asthma, high blood pressure, atrial fibrillation & gout
- Simvastatin, symbicort, bisoprolol fumarate, coumadin & naproxen

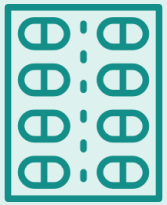
- What side-effects might he expect?
- When should he call his GP?
- What's safe/good for him to eat?

precision.life

Adverse Drug Reactions

2,500
common drugs

15,000
dosage forms



500,000
drug-drug interactions

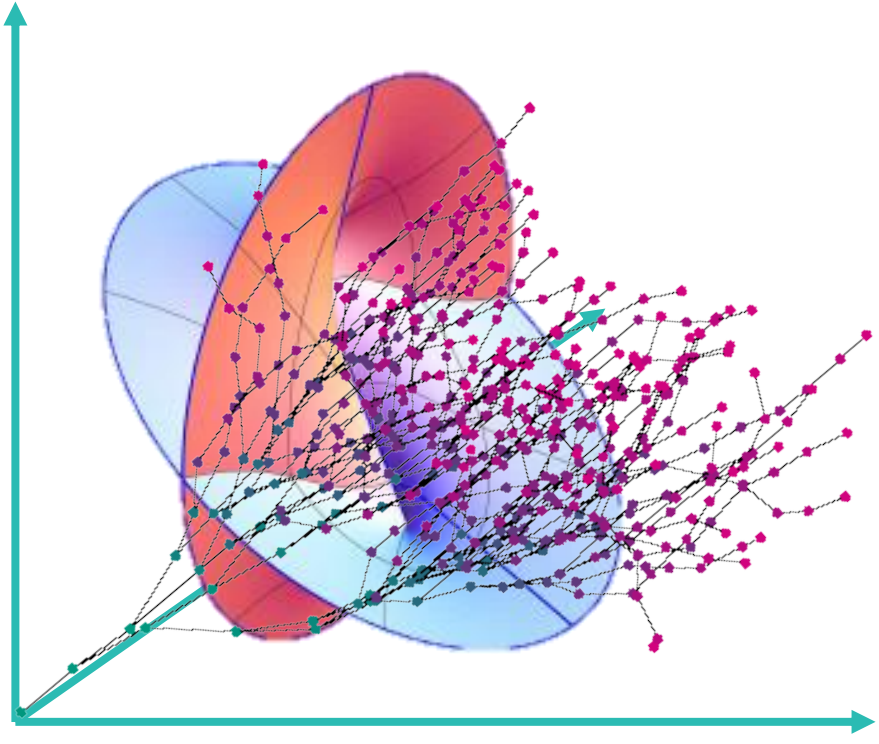
10,000
drug-disease interactions

2,000
drug-food interactions

but... these are just the first-order interactions
our precision.diet API (built on RACE) provides fully
personalized advice considering all combinatorial interactions

precision.life

Analysis of Complex Systems



- Problem space ≈ m^n
  m=no. of states *n*=no. of dimensions
- Easy to get problem spaces of 10^{1000}
- Searching uses short cuts
  - AI/ML, neural nets, GA
- Short cuts miss things and may still require huge CPU/RAM
- RACE Array Logic (tensor algebras) offers provably complete computation quickly using very low CPU/RAM
  - 10^{1000} options > 10 hits in ms

precision.life

RACE Platform

1. Scalable

very large multi-dimensional system models

2. Complete

including all constraints in all dimensions to ensure logical consistency

3. Compact

complete, yet compact representations of complex systems

4. Real-time

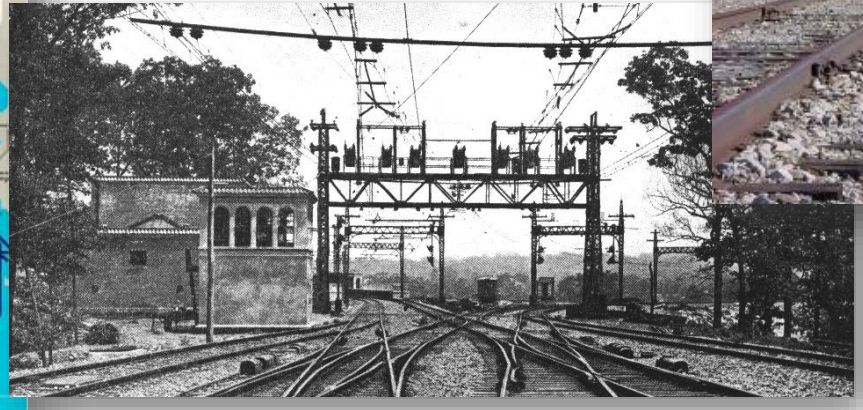
provably complete deduction in real time even on low power devices

precision.life/race
patented (US 6,633,863 / EU 1,062,603) and patents pending

precision.life

Case Study: Danish State Railways

Engineering design and verification



- Verification of railway interlocking systems (track, points, signals…)
- 12,000+ variables and a state space with $>10^{300}$ combinations

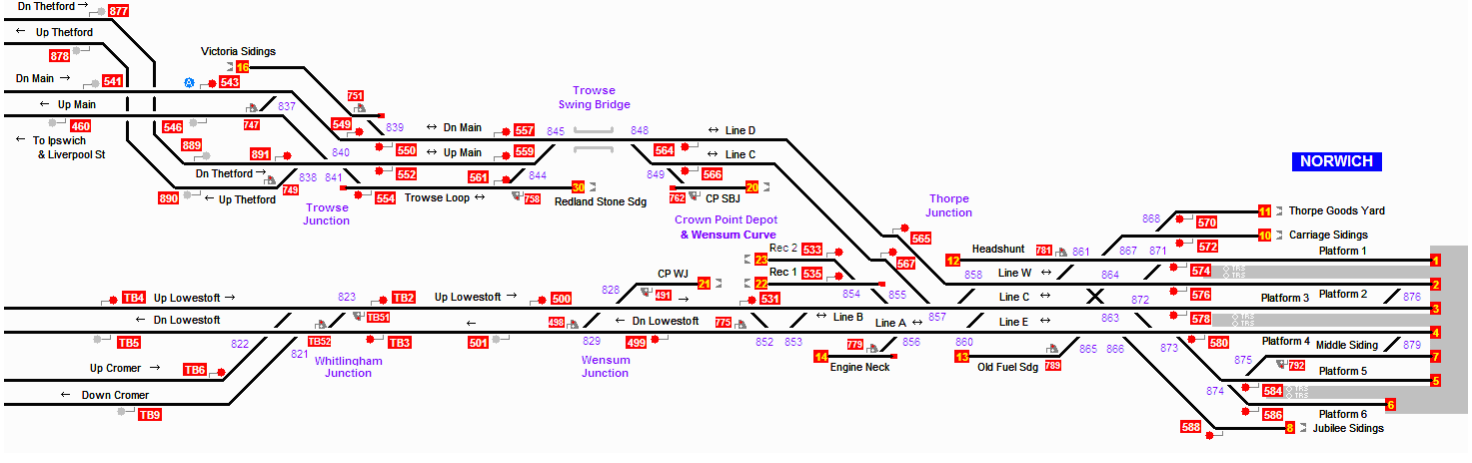
precision.life

Railway Safety Problem

- ALL constraints (physical and logical) must be taken into account to ensure safe and economic operation
- Even small local changes, e.g. a new position of a signal or addition of new points, requires complete validation
- Manual validation of new signal interlocking systems took at least 2 man-years

RACE Solution

- Track topology and connected objects are defined from a CAD tool
- All valid states (i.e. which won't lead to accidents) are determined via constraint resolution, giving a provably complete system state model
  - Entire Danish railway system = 26KB
- Objects functions added to optimize costs and operational efficiency
- 2 man-years validation -> 10ms on a mobile phone



Interactions KGraph/KModel

emc

PubMed

MedlinePlus® Trusted Health Information for You

FDA

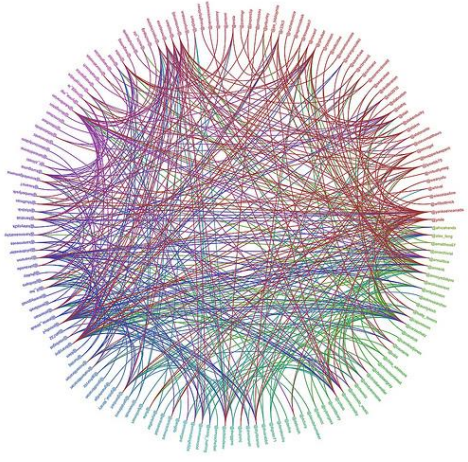
EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

500,000
drug-drug interactions

10,000
drug-disease interactions

2,000
drug-food interactions

Compilation
(20 secs)
RACE Engine



Interactions KModel

precision.life

Interactions KModel

Interactions Knowledge Model

185,000
Food Products
(ingredients, macro/micro nutrients, allergens)

Personal Profile

HealthySwaps App

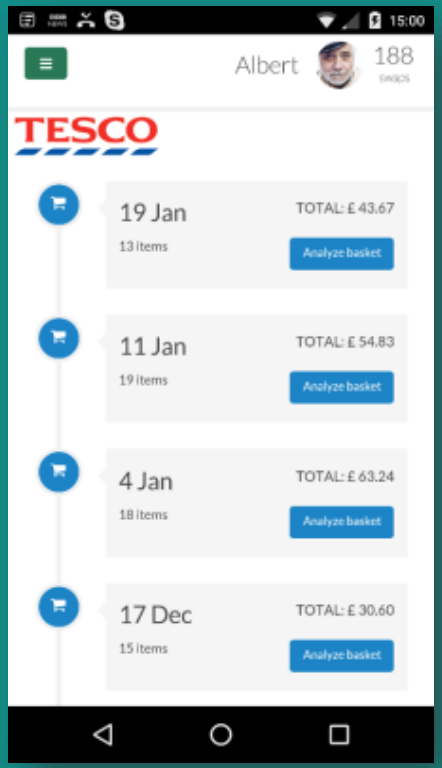
Suggested alternative products that are known to be compatible with the user

Personalized Dietary Advice

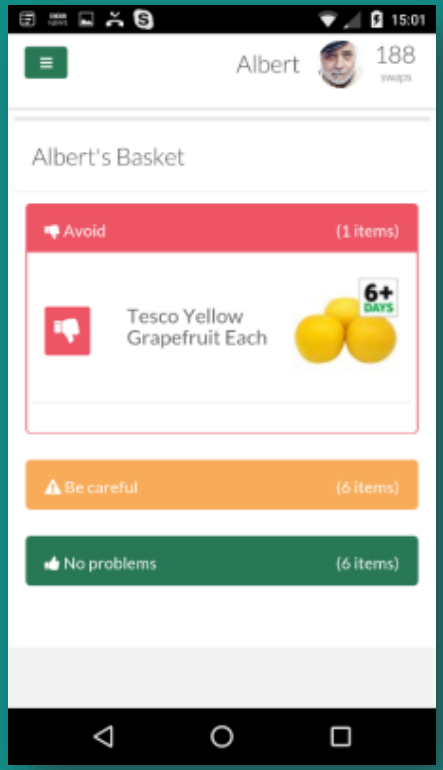
Online Shopping Basket

precision.life

Example – precision.diet



Connect to online shopping basket, use in-store or at home via barcode scanner



Identify food items that are incompatible with your prescriptions, diseases & health goals, and understand risk levels

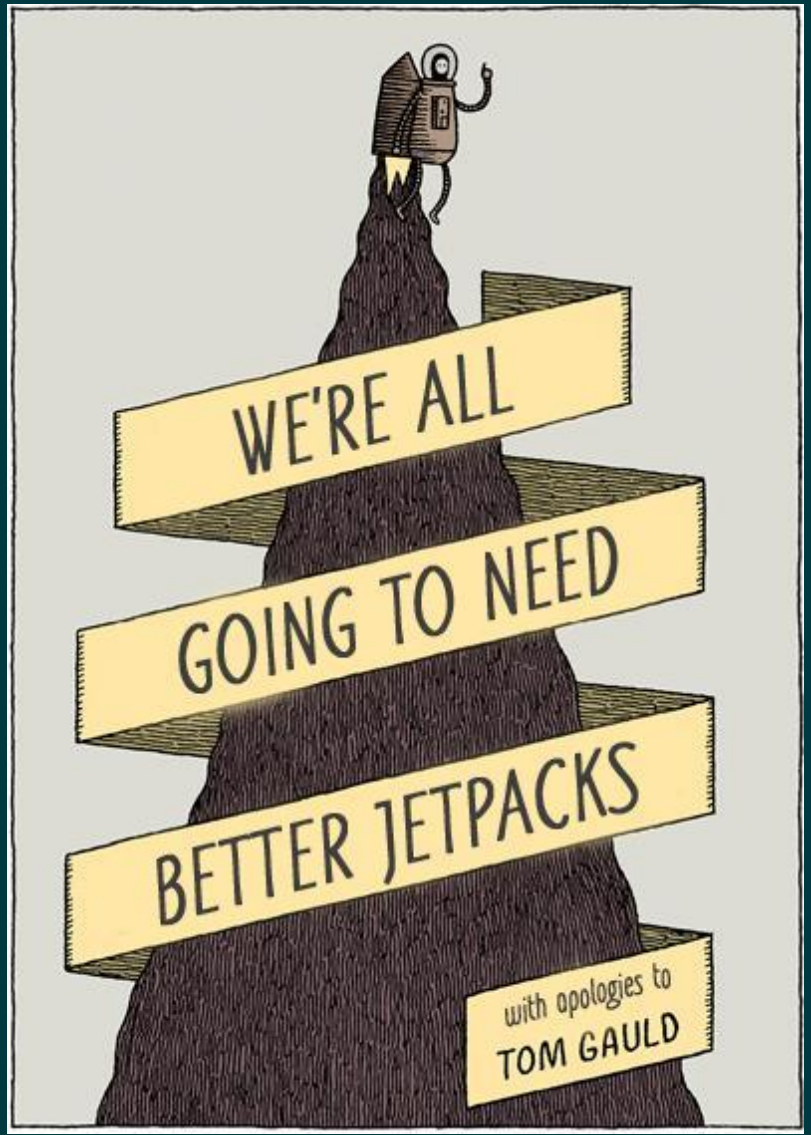
Choose a healthier alternative from same category, all on your own phone with no sharing of your data

Better Tools for Healthcare

- Clinical and patient decision support tools
- At the point of care / in day-to-day life
- Using full power of complex, multi-trait knowledge models

- Improving patient outcomes
- Lowering the cost of care provision

precision.life



Steve Gardner, CEO

E: steve@rowanalytics.com
T: +44 1865 575170 / +44 7799 671644
W: rowanalytics.com

rowanalytics