A Prognostic Atlas of Clinical Medicine:
some initial steps

Berlin Institute of Health Lecture Series
'Frontiers in Translational Medicine –
Scientific and Structural Challenges'

26 November 2021

Harry Hemingway

UCL

NIHR | University College London Hospitals Biomedical Research Centre

HDRUK Health Data Research UK

CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN

What do doctors do? What do hospitals do?

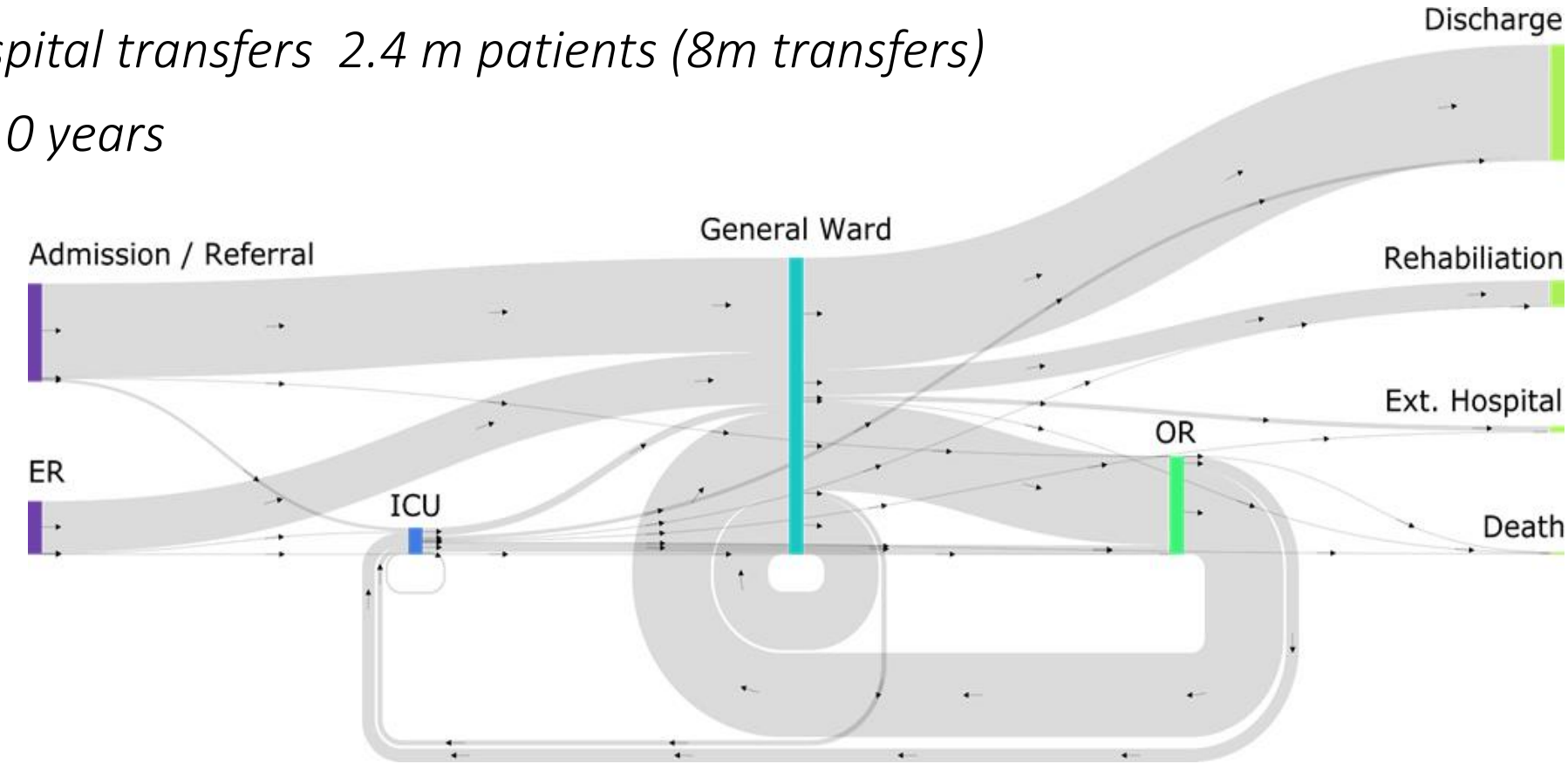
- Conference Facilities
- Interdisciplinary Ward
- Haematology/Oncology
- Gastroenterology
- Rheumatology
- Neurology
- Neurosurgery
- Orthopaedics
- ENT
- Audiology/Phoniatrics
- Cardiology
- Surgery
- Mothers and Newborns
- Neonatology
- Dialysis

- Haematology/Oncology
- Gastroenterology/Endocrinology
- Nephrology
- Neurology/Stroke Unit
- Neurosurgery
- Orthopaedics
- Orthopaedics
- Urology
- Cardiology
- Chirurgic
- Clinical Directorates
- Delivery Suite
- Gynaecology
- Clinical Directorates

Doctors treat patients and they keep records > EHR data

In-hospital transfers 2.4 m patients (8m transfers)

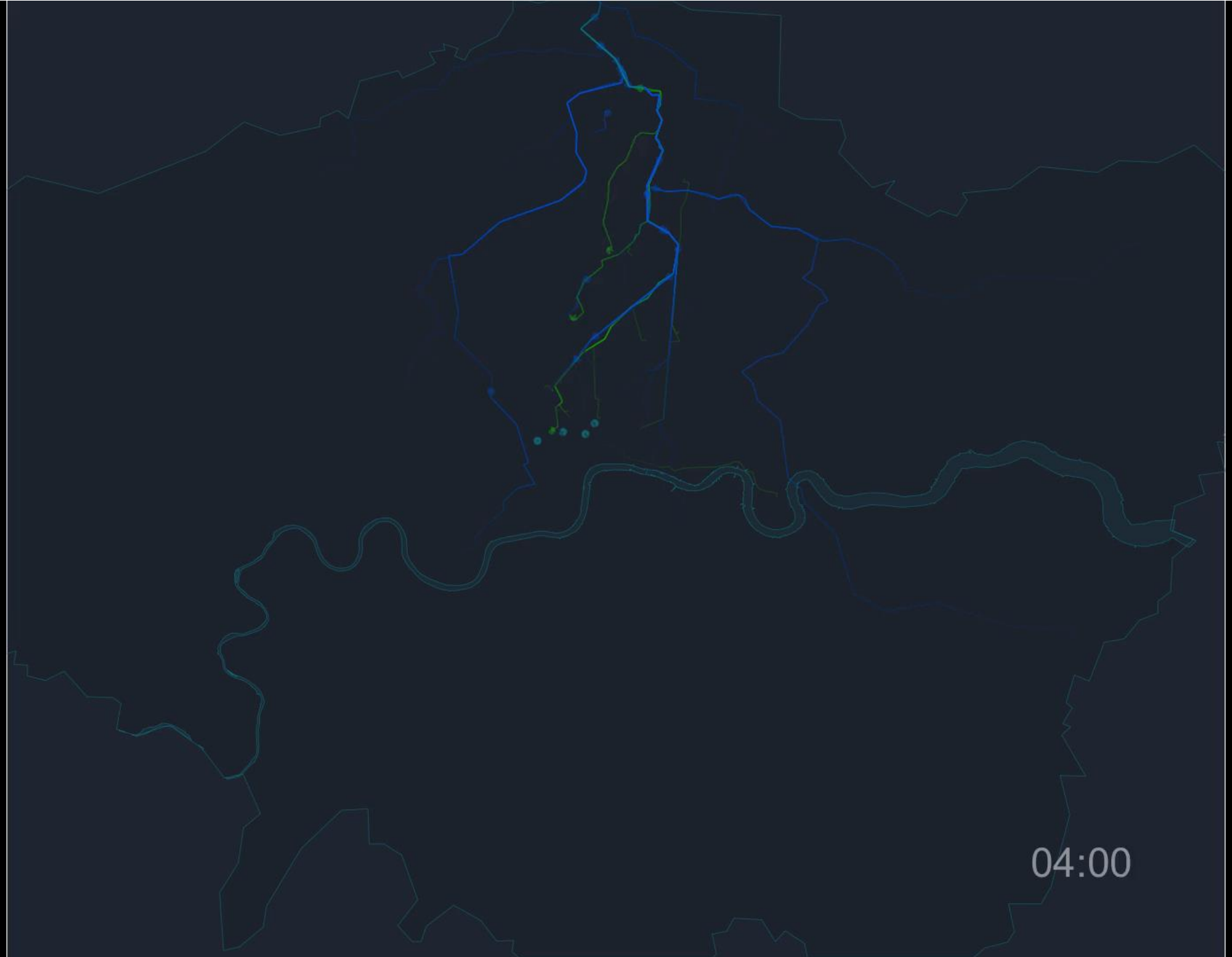
over 10 years



Prof Felix Balzer

CHARITÉ
UNIVERSITÄTSMEDIZIN BERLIN

Institute of Medical Informatics, Health Data Platform

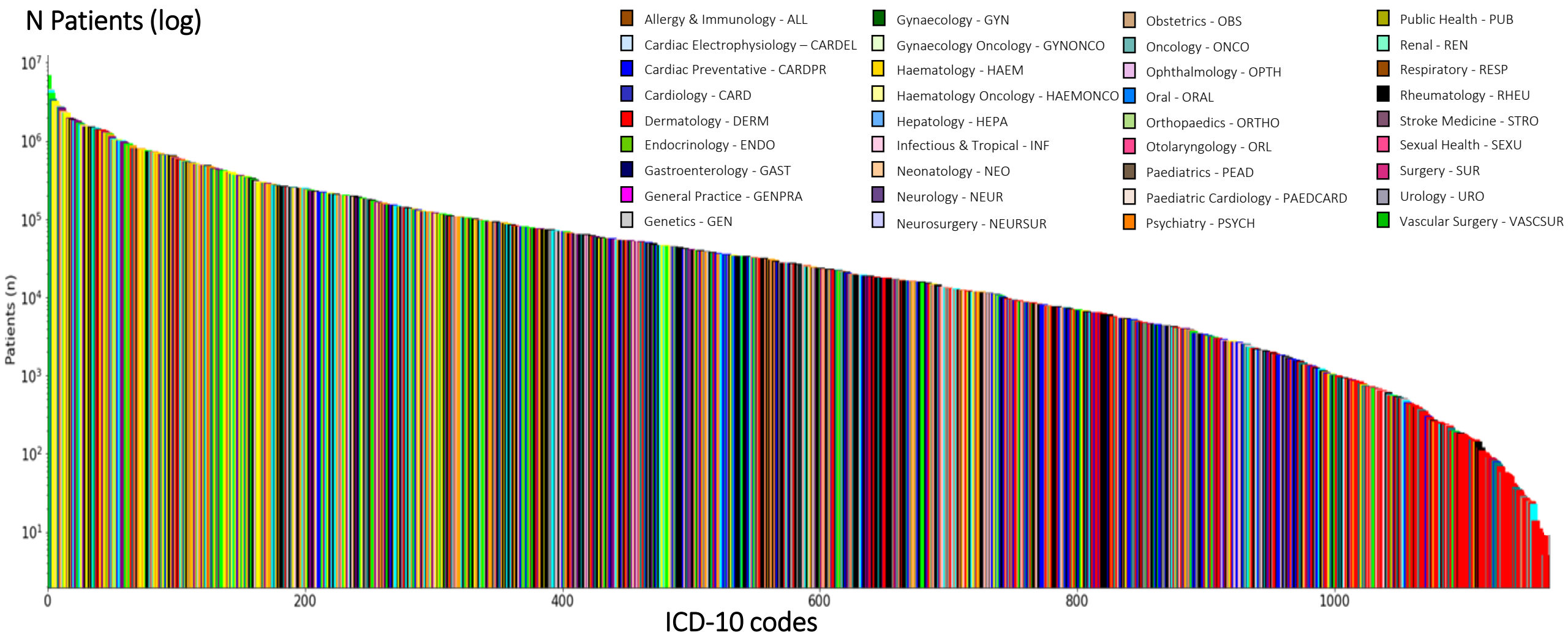

04:00

Clinicians diagnose diseases and write it down

56 million people in England alive at 30 Jan 2020 (the day before COVID case #1) and updated hospitalisations, currently 250 million

Prof Cathie Sudlow OBE

Prof Spiros Denaxas



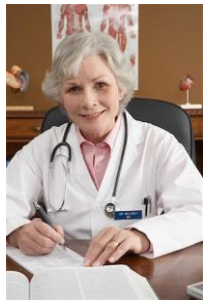
Healthcare systems across the world know remarkably little about which patients have which diseases, in which combinations, and with what outcomes.

International unmet need

Lack of findable, accessible, useful personalised risk for people with each disease

'People like me'

'Patients like mine'

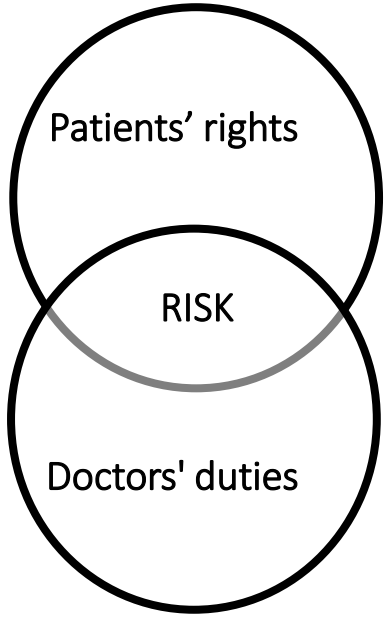


Existing risk information:

- Low coverage of: <5% of diseases have a prognostic model
- Not 'for' patients
- Mono morbid approach (but patients commonly multimorbid)
- Serendipitous ‘one disease at a time’ generation of prognostic models: no systematic framework

Prognosis and medical ethics

"Patients have the right
to discuss risks"



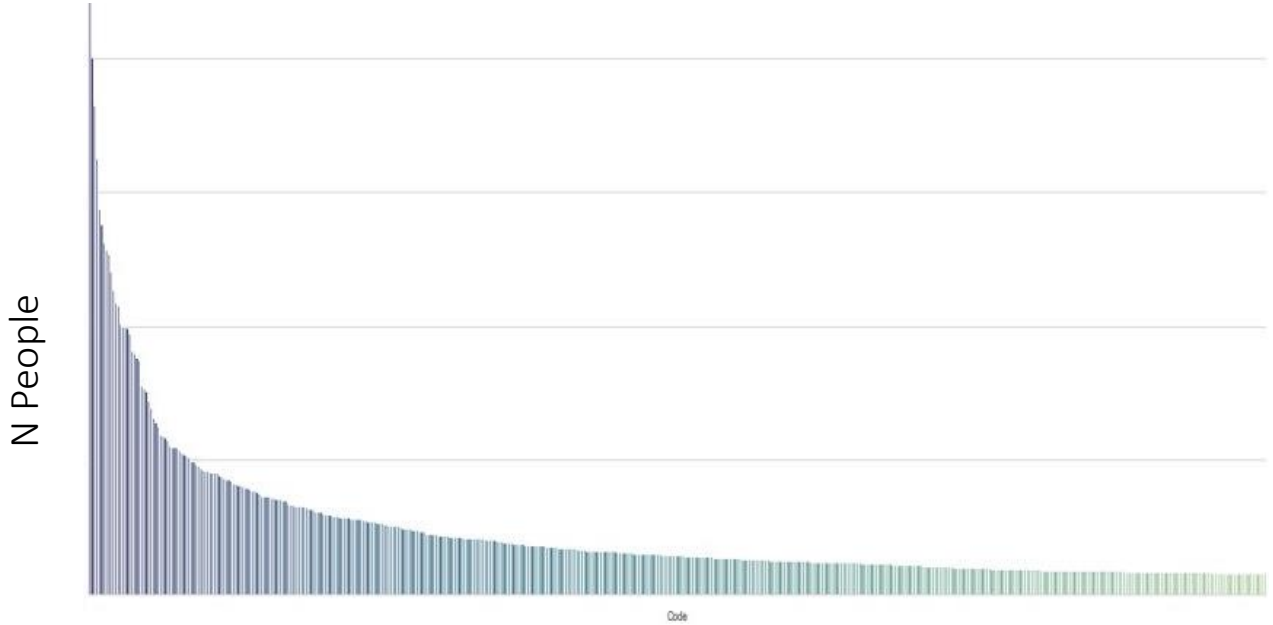
"You must …
share information on
likely progression"

AMA AMERICAN MEDICAL ASSOCIATION

General Medical Council

Can we be more systematic and do this for each and every disease?

Democratic approaches to answering ubiquitous questions across the ‘long tail’ of clinical medicine



All specialities ↔ All clinicians ↔ All diseases ↔ All people with diseases

Prognosis: massive need for more, and more actionable, research

BMJ

Prognosis research strategy (PROGRESS) 1: A framework for researching clinical outcomes.
Hemingway H, Croft P, Perel P, Hayden JA, Abrams K, Timmis A, Briggs A, Udumyan R, Moons KGM, Steyerberg EW, Roberts I, Schroter S, Altman DG, Riley RD. *BMJ.* 2013 Feb 5;346:e5595.

PLOS | MEDICINE

Prognosis Research Strategy (PROGRESS) 2: Prognostic Factor Research.
Riley RD, Hayden JA, Steyerberg EW, Moons KG, Abrams K, Kyzas PA, Malats N, Briggs A, Schroter S, Altman DG, Hemingway H. *PLoS Med.* 2013 Feb;10(2):e1001380.

PLOS | MEDICINE

Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, Riley RD, Hemingway H, Altman DG. *PLoS Med.* 2013 Feb; 10(2):e1001381.

BMJ

Prognosis Research Strategy (PROGRESS) 4: Stratified Medicine Research. Hingorani AD, Windt DA, Riley RD, Abrams K, Moons KG, Steyerberg EW, Schroter S, Sauerbrei W, Altman DG, Hemingway H. *BMJ.* 2013 Feb 5;346:e5793.

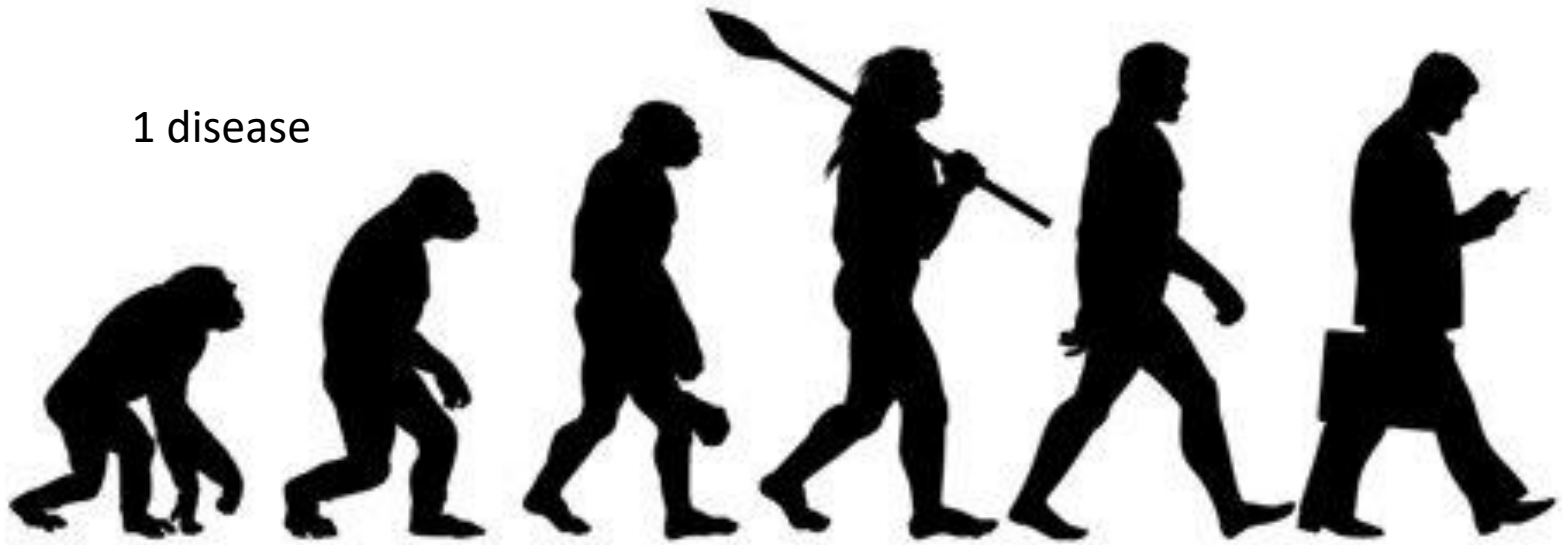
Approach: from serendipity to systems

All diseases

100s of diseases

12 diseases

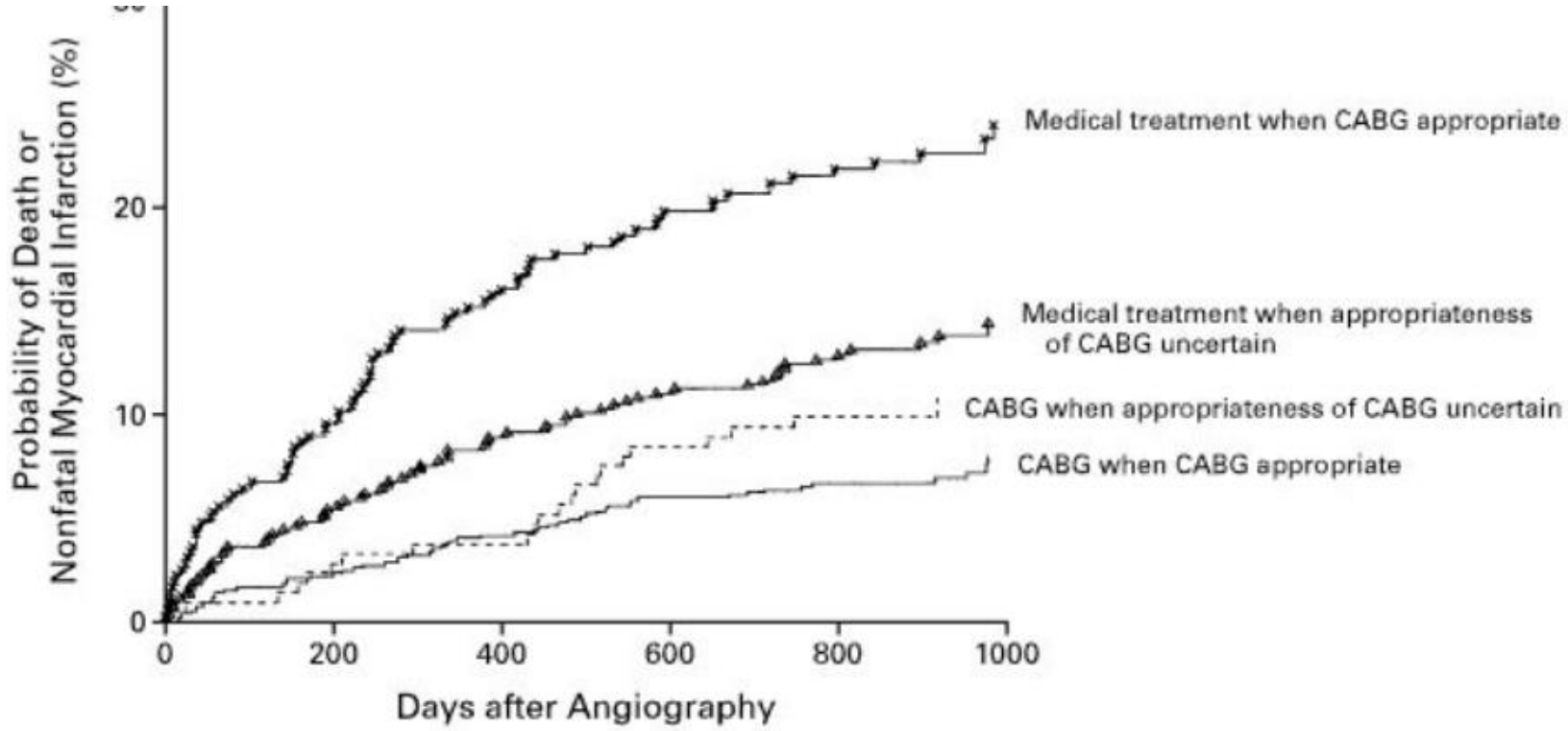
1 disease



One disease at a time

1 Disease: Angiographic Coronary Artery disease:

Clinical data: 5 hospitals, ‘weighing paper case notes’
Outcomes from hospital EHR linkage (first time in England)

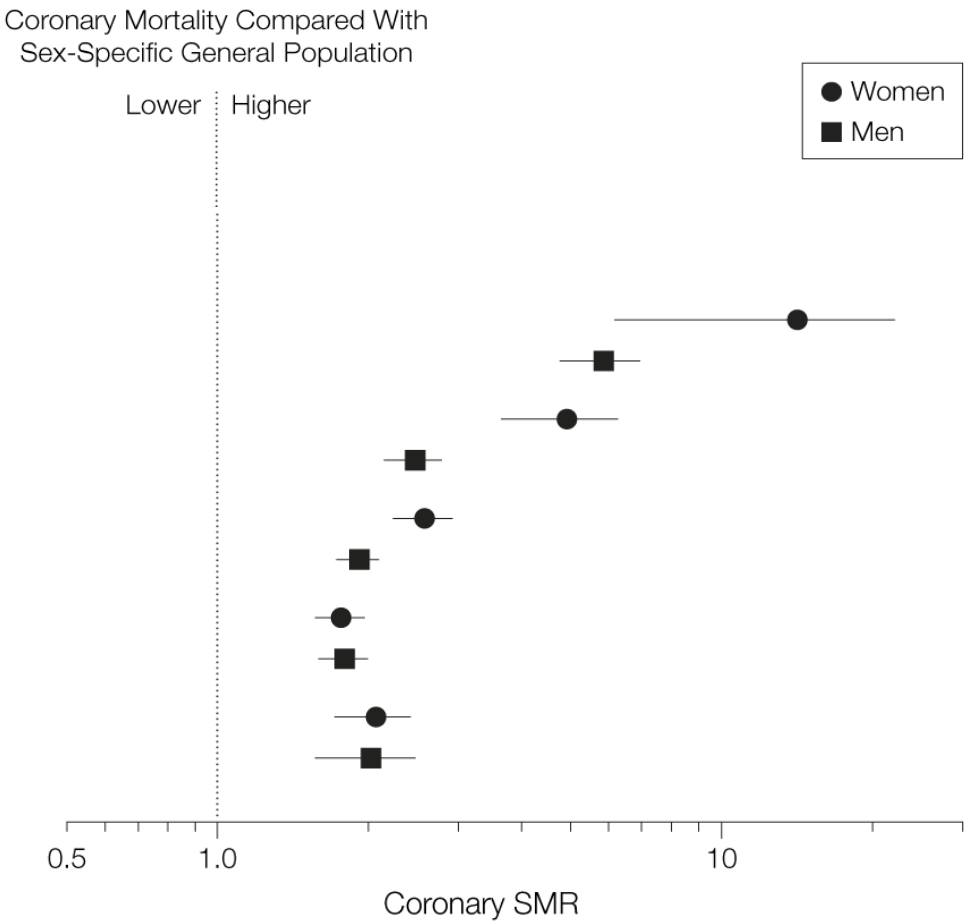


Hemingway et al. NEJM 2001

1 Disease: angina

Clinical data: whole country nationwide in Finland
CD sent in the post!

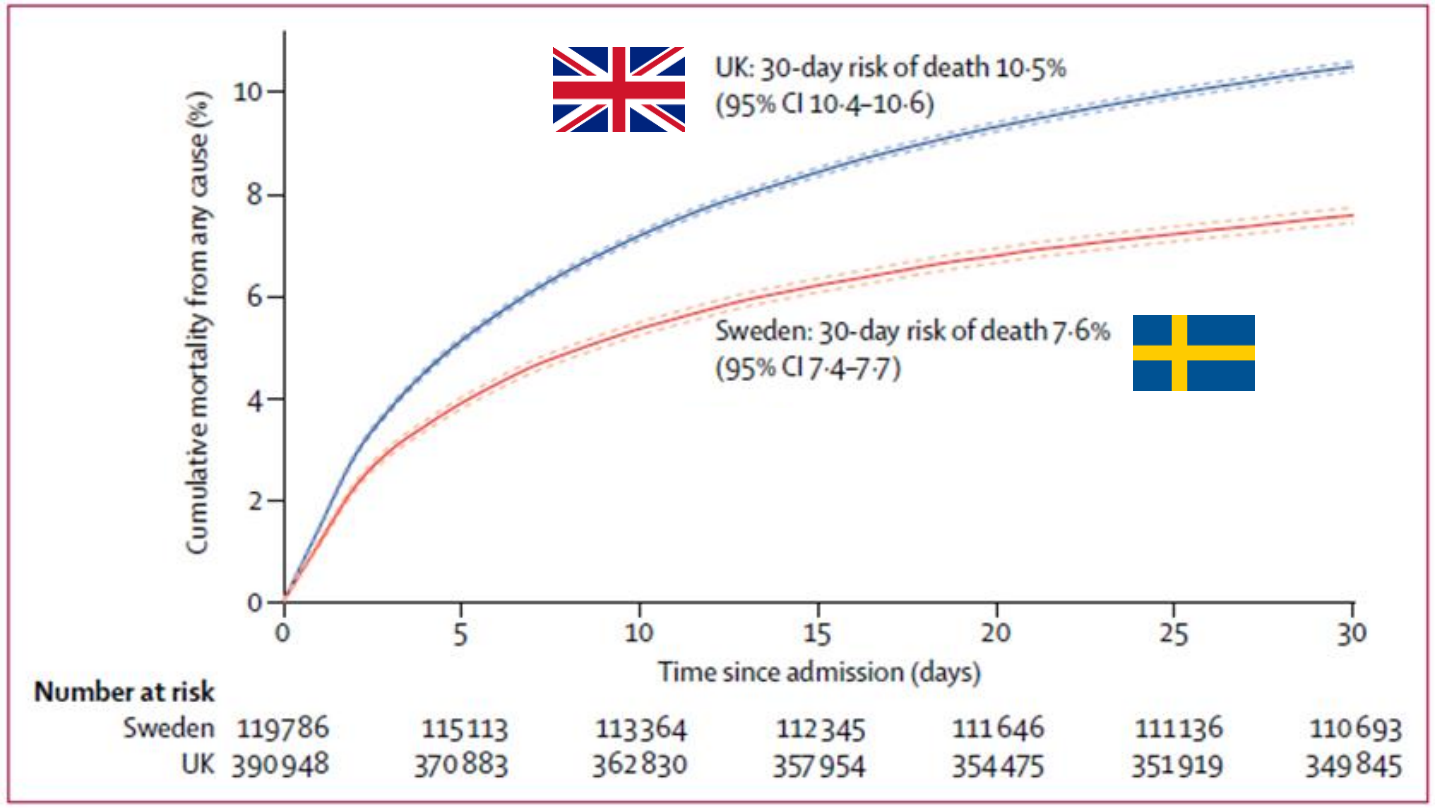
Age Group, y	Sex	Observed/ Expected Deaths	SMR (95% CI)	*P* Value for Interaction
		Test-Positive Angina		
45-54	Women	11/1	12.1 (6.06-21.7)	.02
	Men	120/21	5.63 (4.71-6.74)	
55-64	Women	55/12	4.69 (3.60-6.11)	<.001
	Men	229/95	2.40 (2.11-2.73)	
65-74	Women	221/88	2.50 (2.20-2.86)	<.001
	Men	452/242	1.87 (1.70-2.05)	
75-84	Women	322/187	1.72 (1.54-1.92)	.83
	Men	329/188	1.75 (1.57-1.95)	
85-89	Women	127/64	2.00 (1.68-2.37)	.87
	Men	74/39	1.93 (1.54-2.42)	



Hemingway et al. JAMA 2006

1 Disease: acute myocardial infarction

Whole country nationwide in UK and Sweden
Disease registry: manual data entry, not part of EHR



Acute myocardial infarction: a comparison of short-term survival in national outcome registries in Sweden and the UK

Sheng-Chia Chung, Rolf Gedeborg, Owen Nicholas, Stefan James, Anders Jeppsson, Charles Wolfe, Peter Heuschmann, Lars Wallentin, John Deanfield, Adam Timmis, Tomas Jernberg, Harry Hemingway

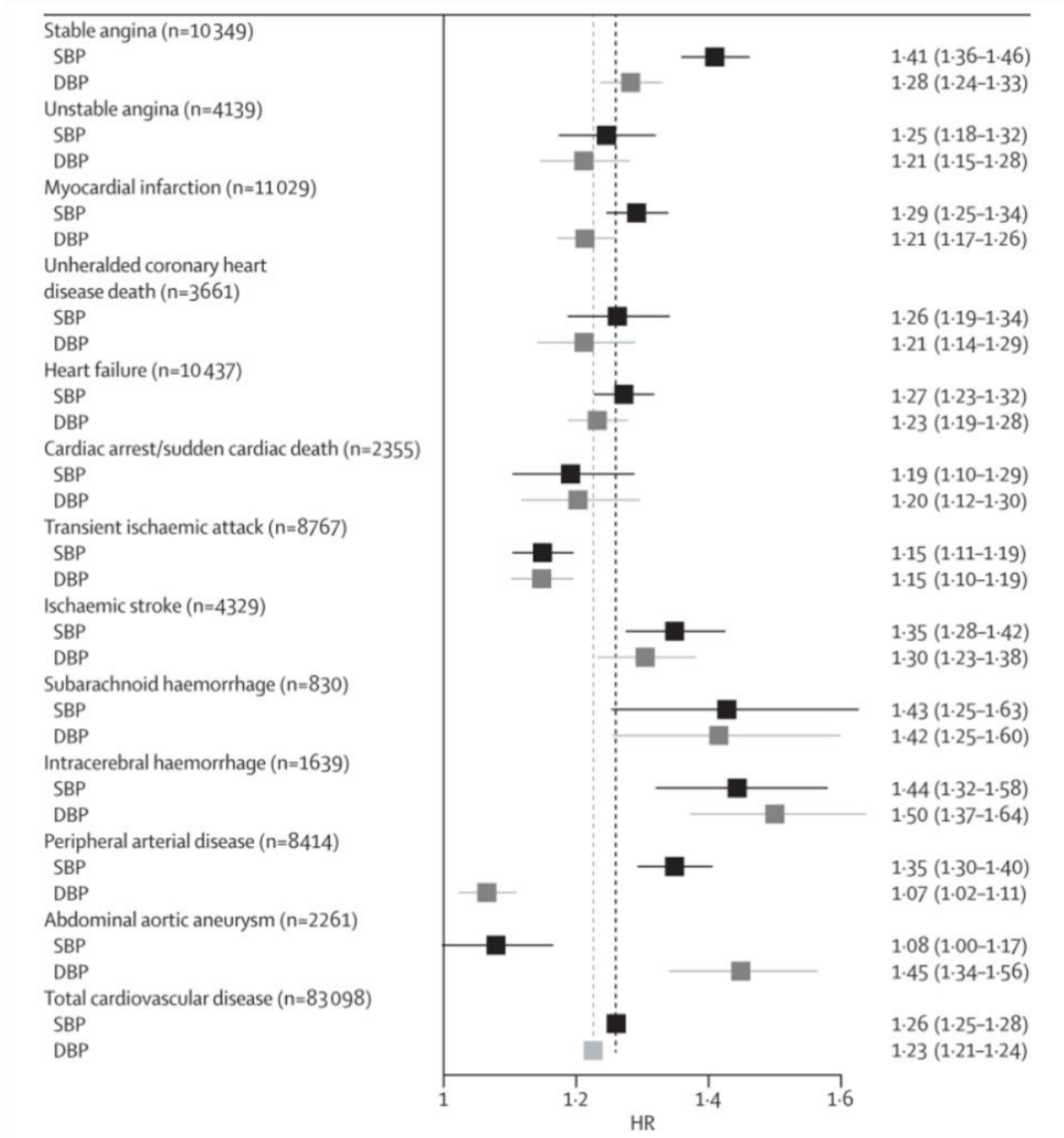
12 diseases at a time

12 diseases: Higher resolution ‘CVD’

3% sample of England (1.2m people)
EHR (primary care linked to secondary care)
clinically recorded blood pressure

Informed American and European guidelines

Risks differ across diseases

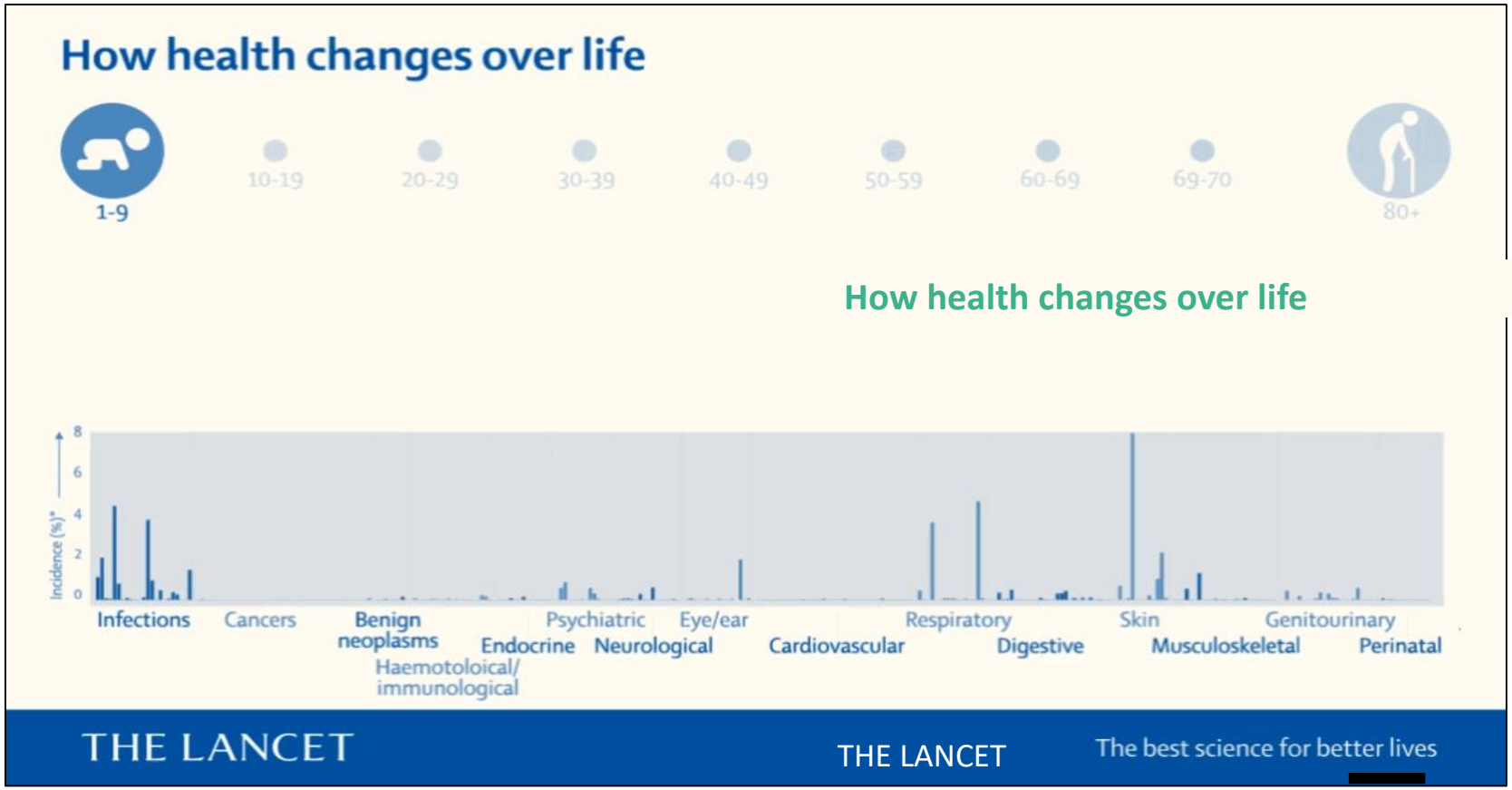


Rapsomaniki et al. The Lancet 2014

Many common diseases

A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service

Valerie Kuan, Spiros Denaxas, Arturo Gonzalez-Izquierdo, Kenan Direk, Osman Bhatti, Shanaz Husain, Shailen Sutaria, Melanie Hingorani, Dorothea Nitsch, Constantinos A Parisinos, R Thomas Lumbers, Rohini Mathur, Reecha Sofat, Juan P Casas, Ian C K Wong, Harry Hemingway, Aroon D Hingorani



How health changes over life

Dr Valerie Kuan

EHR phenotyping

1001, 2000-01-01, 23,1,NULL,I48
1121, 2013-05-04, 7,1,3,5,14AN.00
1121, 2011-05-21, 81,1,9, G573100
1511, 1993-01-11, 91,1,6,9hF1.00
1511, 199-03-11, 91,1,6, G573100
9913, 2012-05-21, 81,1,9, G573100
67222, 1994-11-01,1234,1,3,7L1H300
1001, 1994-08-11,1234,1,3,7L1H300
1001, 1993-01-01, 253,1,1,793Mz00
1231, 2012-03-03, 23,1,123,K65
1121, 2013-05-04, 7,1,3,5,14AN.00
1121, 2011-05-21, 81,1,9, G573100
1511, 1993-01-11, 91,1,6,9hF1.00
1511, 199-03-11, 91,1,6, G573100
9913, 2012-05-21, 81,1,9, G573100
67222, 1994-11-01,1234,1,3,7L1H300
67222, 1995-12-21,1234,1,3,7L1H300
67222, 1991-03-03,1234,1,3,7L1H310
682444, 1993-01-01, 253,1,1,793Mz00
1121, 2013-05-04, 7,1,3,5,14AN.00
1121, 2011-05-21, 81,1,9, G573100
1511, 1993-01-11, 91,1,6,9hF1.00
1511, 199-03-11, 91,1,6, G573100
67222, 1995-12-21,1234,1,3,7L1H300
67222, 1994-11-01,1234,1,3,7L1H300

EHR Phenotyping

Disease status
Severity
Onset
Certainty

Open EHR Phenotyping library

HDRUK Phenotype Library

Home Phenotypes Concepts API About Log in

The HDR UK Phenotype Library is a comprehensive, open access resource providing the research community with information, tools and phenotyping algorithms for UK electronic health records.

Search our Phenotype library

753	1462	101562	22	14
Phenotypes	Concepts	Clinical Codes	Data Sources	Coding Systems

A Reference Catalogue of Human Diseases

Connected. The Phenotype Library is accessible via an API to support interoperability, is integrated with health dataset information in HDR-UK's Innovation Gateway, and hosts content from numerous contributing organisations.

Patient-focused. The Library is enabling important research to improve patient health and well-being. Content spans major disease areas, including heart disease, cancer, COVID-19 and other common and rare human health conditions. Curated collections from contributors such as the HDR UK BREATHE Hub for respiratory health share clinical expertise to tackle critical research questions.

Cutting-edge. Built with a focus on computability, this resource aims to drive the next generation of research methods. Integration with Phenoflow enables executable implementations of the phenotypes in our collection, while the API and R package client facilitate integration of the Library content directly into other analysis workflows.

https://phenotypes.healthdatagateway.org

Export Print

CSV
JSON
XML

Asthma

Eleanor L Axson, Jennifer K Quint

Type	Disease or Syndrome
ID	PH12
Version ID	24
Data Sources	Clinical Practice Research Datalink GOLD , Clinical Practice Research Datalink AURUM , Hospital Episode Statistics APC for CPRD GOLD , Hospital Episode Statistics APC for CPRD Aurum , Death Registration data for CPRD GOLD , Death Registration data for CPRD Aurum , UK Biobank
Valid event data range	01/01/2001 - 31/12/2019
Sex	Female, Male
Agreement Date	2020-06-03
Coding system	Read codes v2 ICD10 codes SNOMED codes UKBioBank codes ICD11 codes
Tags	BREATHE Phenotype Library

Definition

These codes will capture asthma ever, not just current asthma. These codes are not intended to be mandatory, but are to be used as a starting point for the identification of asthma in routine EHR. Each study may differ in the sensitivity and specificity of the coding required.

For those interested in further discrimination of asthma phenotypes, we refer you to Nissen et al. 2019.

F. Nissen, Douglas, I. J., Mullerova, H., Pearce, N., Bloom, C. I., Smeeth, L., and Quint, J. K., ?Clinical profile of predefined asthma phenotypes in a large cohort of UK primary care patients (Clinical Practice Research Datalink)?, J Asthma Allergy, vol. 12, pp. 7-19, 2019.

Validation of Read Codes for the Identification of COPD in CPRD

Quint et al. validated a set of Read codes for the identification of COPD in CPRD in 2014. Using diagnostic codes alone, the positive predictive value (PPV) was 86.5% (77.5?92.3%). Requiring a diagnostic code, spirometry measures, and specific medication increased PPV to 89.4% (80.7?94.5%) but reduced case numbers by 10%.

Validations of EHR Phenotypes

- Specialist adjudication *vs* standard
- Clinical relevance e.g. mapping to guidelines + quality initiatives
- Outcomes and Prognosis
- Genetic / molecular / aetiology
- Concordance across settings (1ry, 2ry, 3ry)
- Transportability across health systems and nations

Denaxas et al. JAMIA 2018

Uses and insights from EHR Phenotypes

- Specialist adjudication *vs* standard
- Clinical relevance e.g. mapping to guidelines + quality initiatives
- Outcomes and Prognosis
- Genetic / molecular / aetiology
- Concordance across settings (1ry, 2ry, 3ry)
- Transportability across health systems and nations

Denaxas et al. JAMIA 2018

Use of 100s of EHR disease phenotypes: pre-prototype of prognostic atlas for pandemic response

Pre-print 21 March 2021
The day before UK Lockdown 1

Estimating excess 1-year mortality associated with the COVID-19 pandemic according to underlying conditions and age: a population-based cohort study

Amitava Banerjee, Laura Pasea, Steve Harris, Arturo Gonzalez-Izquierdo, Ana Torralbo, Laura Shallcross, Mahdad Noursadeghi, Deenan Pillay, Neil Sebire, Chris Holmes, Christina Pagel, Wai Keong Wong, Claudia Langenberg, Bryan Williams, Spiros Denaxas, Harry Hemingway

OurRisk.CoV

Prof Ami Banerjee

OurRisk.CoV

Underlying conditions or ongoing treatments

Heart failure

More than one of the listed conditions?

(•) Yes ( ) No

Age (years)

60-70

Sex

(•) Man ( ) Woman

Calculate risk

Adjust our risk

- Welcome to OurRisk.CoV
- Who is OurRisk.CoV for?
- How to use OurRisk.CoV
- What should I do with this information?
- Tell us what you think

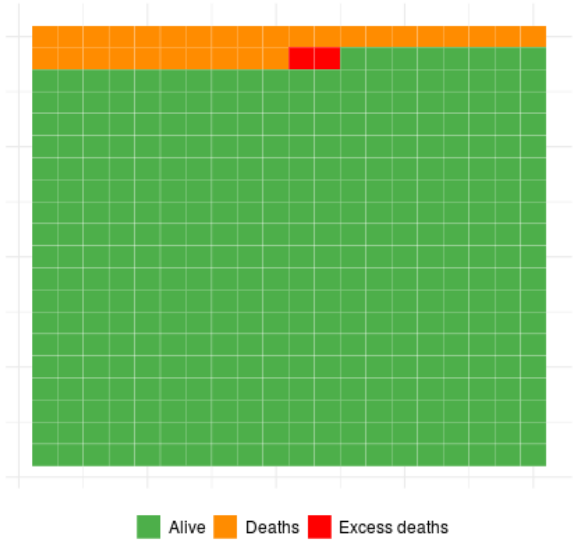
Risk calculator | Further information

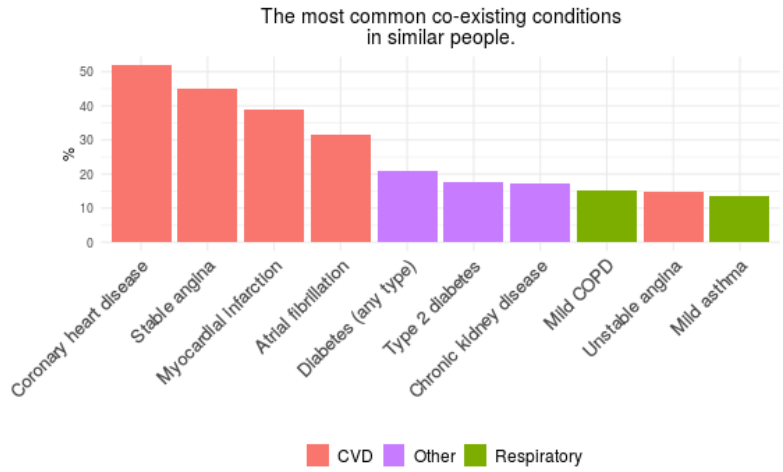
In similar people in England:

1-year risk of death

Baseline	Projected during the pandemic
7.4% (6.6 - 8.2%)	7.9% (7 - 8.7%)

Expected deaths amongst individuals with similar characteristics





http://covid19-phenomics.org/PrototypeOurRiskCoV.html

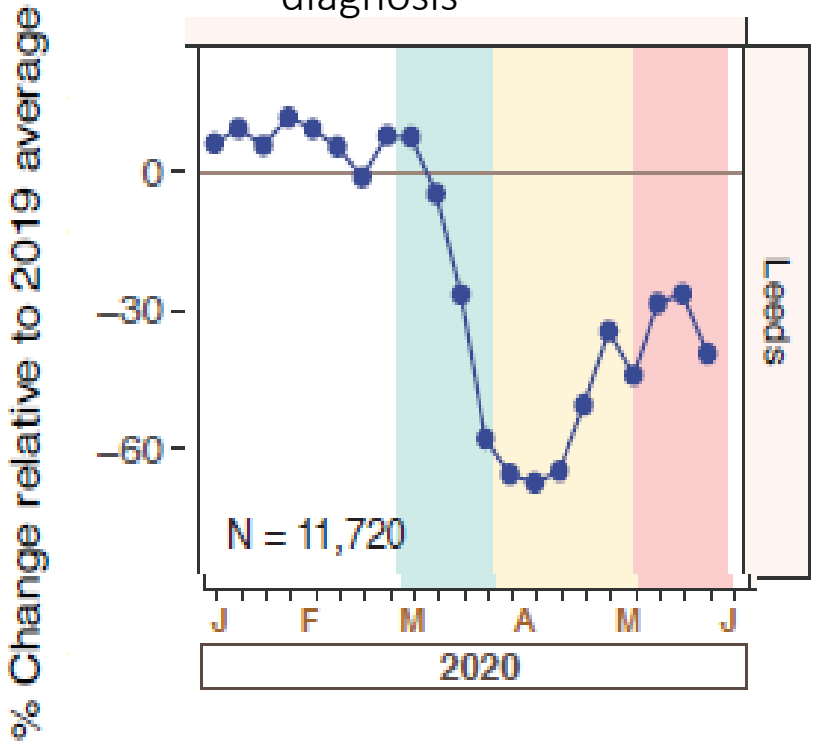
"At last I have some information *about me that I can act on*"

1.4m page views

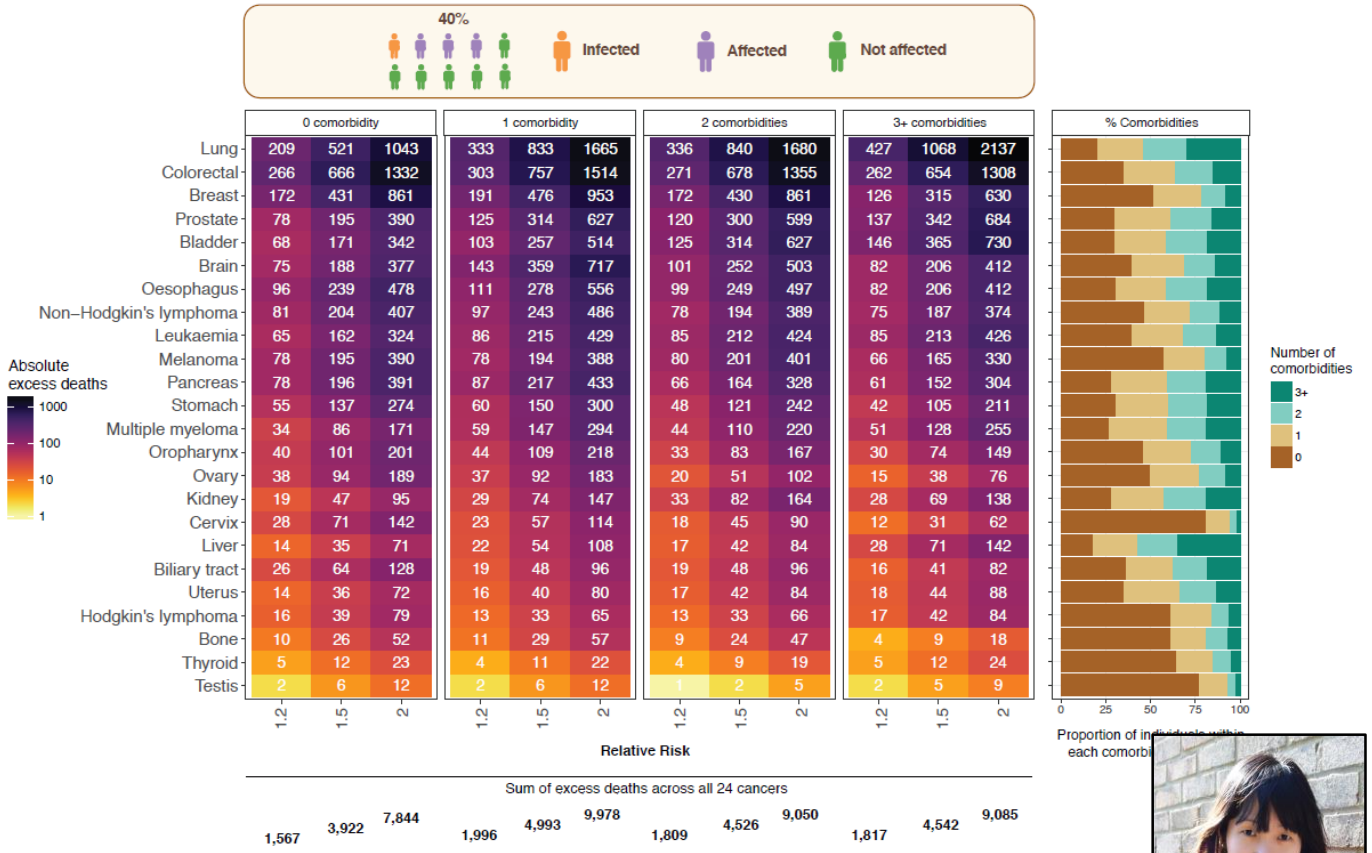
Dr Laura Pasea

Use of 100s of EHR disease phenotypes: pre-prototype of prognostic atlas lockdown 1 - cancer

Referrals for urgent cancer diagnosis

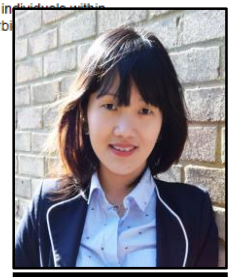


Atlas as ‘canary in the mine’



Estimating excess mortality in people with cancer and multimorbidity in the COVID-19 emergency

Alvina G. Lai, PhD.[1,2✉], Laura Pasea, PhD.[1,2*], Amitava Banerjee, DPhil[1,2,3*], Spiros Denaxas, PhD.[1,2,6,7], Michail Katsoulis, PhD.[1,2], Wai Hoong Chang, MSc[1,2], Bryan Williams, PhD.[4,5,6], Deenan Pillay, PhD.[8], Mahdad Noursadeghi, PhD.[8], David Linch, FMedSci[8,9], Dennilynn Hughes, FRCPath[10,11], Martin D. Forster, PhD.[4,10], Clare Turnbull, PhD.[12], Natalie K. Fitzpatrick, MSc[1,2], Kathryn Boyd, MD[13], Graham R. Foster, PhD.[14], DATA-CAN[15], Matt Cooper, PhD.[15], Monica Jones, PGDip[15], Kathy

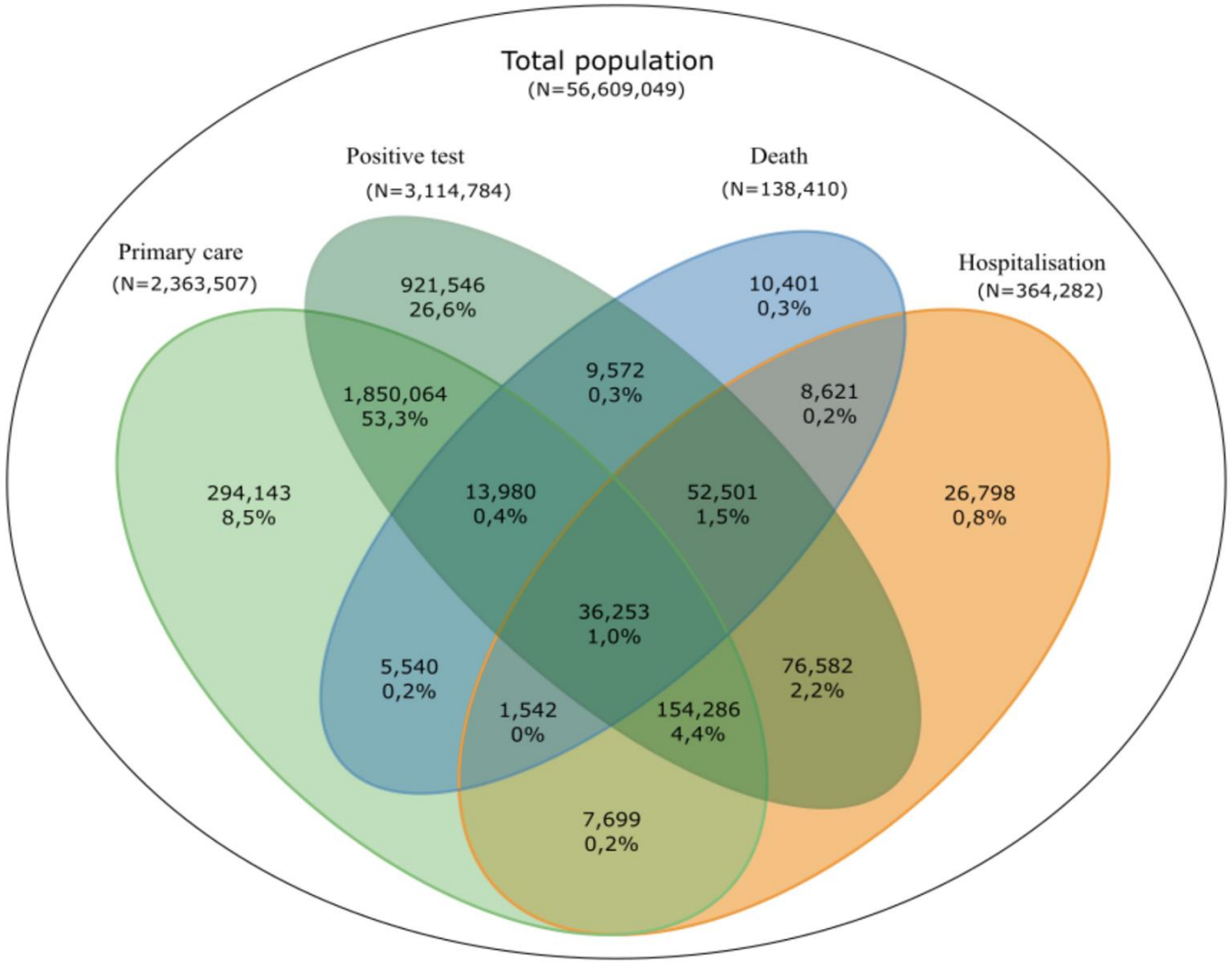


Dr Alvina Lai

SCAN FOR PAPER

.....subsequent NHS Recovery Plan in July 2020 prioritizes cancer services

2021- a step up in the scale and depth of data in England accessible by researchers



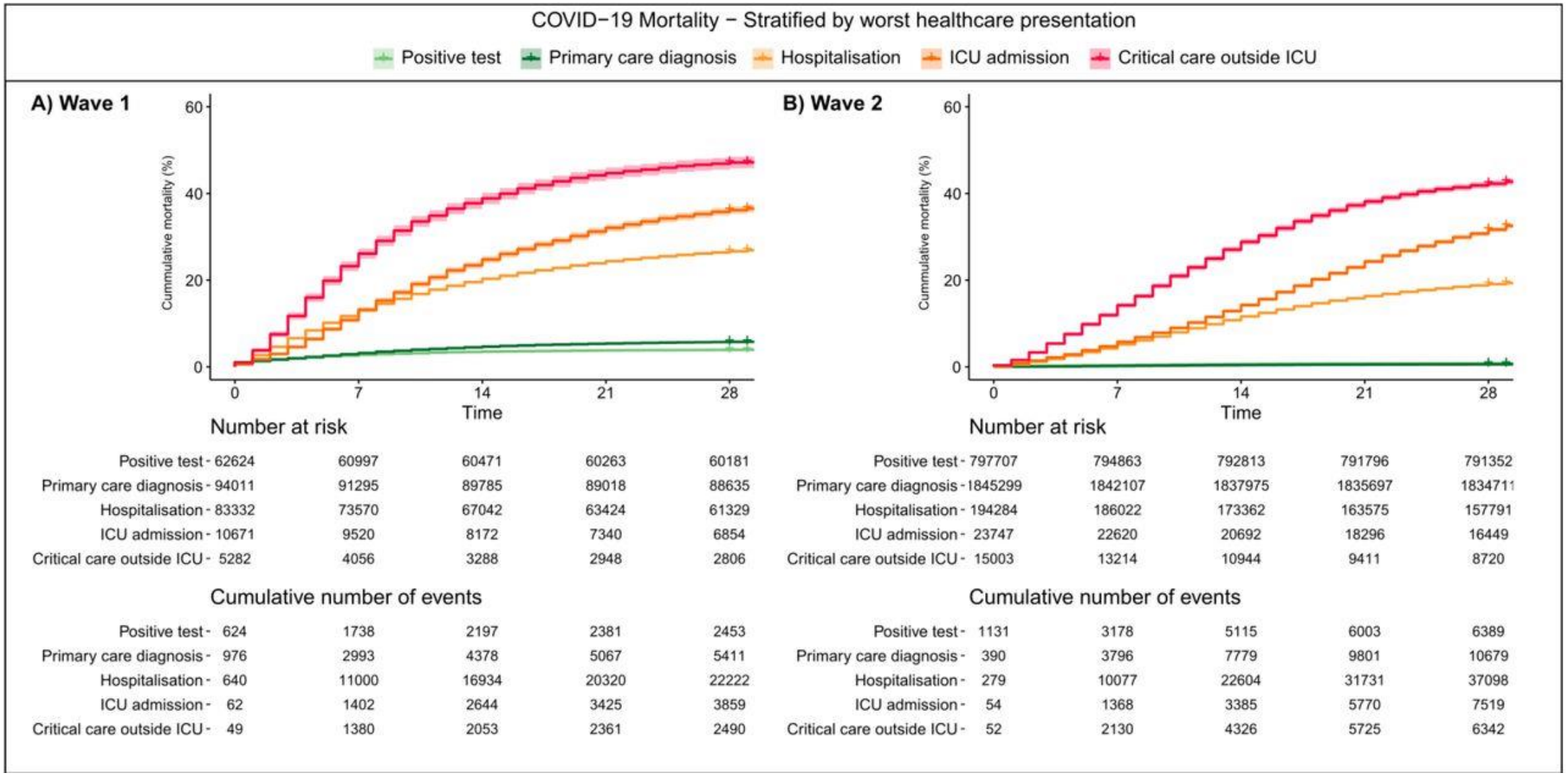
Prof Cathie Sudlow OBE

Dr Johan Thygesen

Dr Chris Tomlinson

Thygesen et al. medRxiv 9 November 2021

COVID trajectories in 56m people



Thygesen et al. medRxiv (preprint) 9 November 2021

All diseases

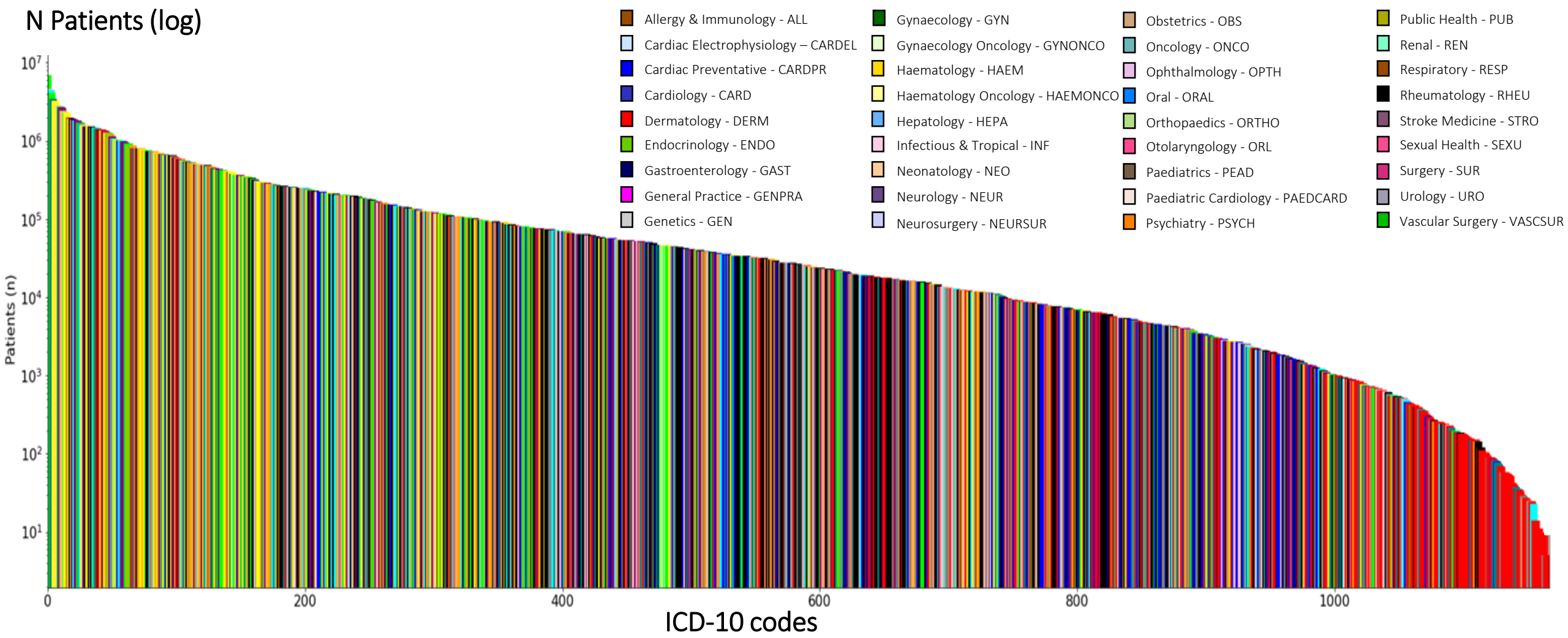
across common-rare disease continuum
that leave a digital trace

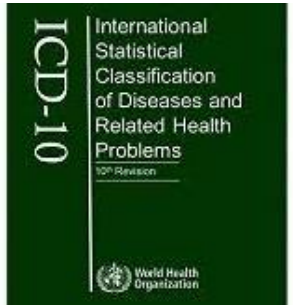
Clinicians diagnose diseases and write it down

56 million people in England alive at 30 Jan 2020 (the day before COVID case #1) and updated hospitalisations, currently 250 million

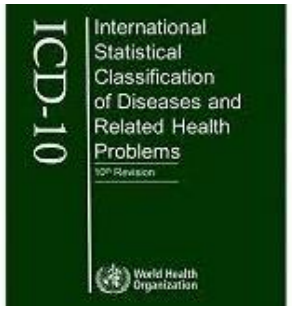
Prof Cathie Sudlow OBE

Prof Spiros Denaxas





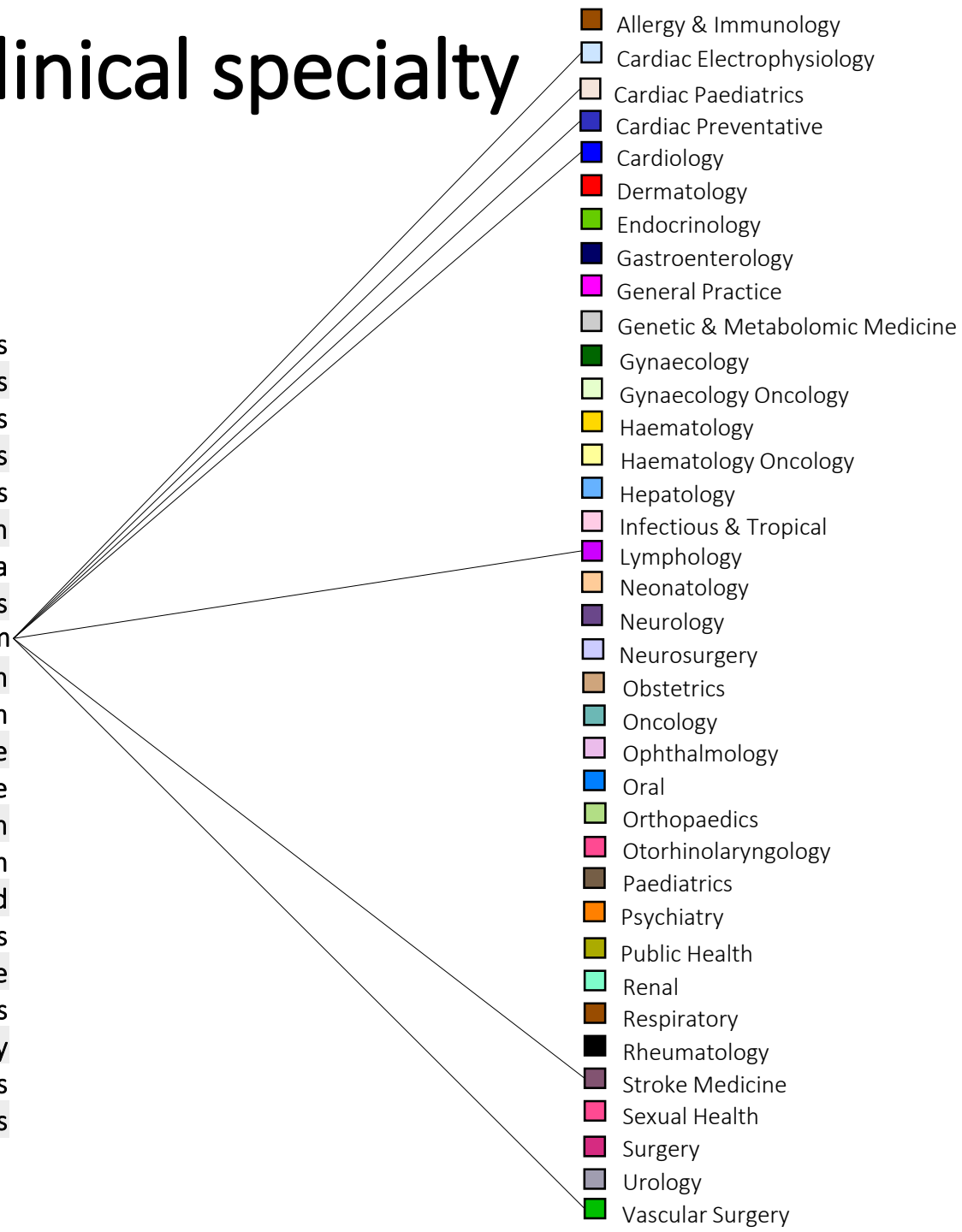
- It is the *most* widely used terminology internationally in health systems, and research (morbidity and mortality)
- 22 chapters (based on pathology, or physiology or anatomy
- 12000 of 17 000 unique ICD-10 4 character codes *are used in practice*
- Hi fidelity across common and rare disease abundance



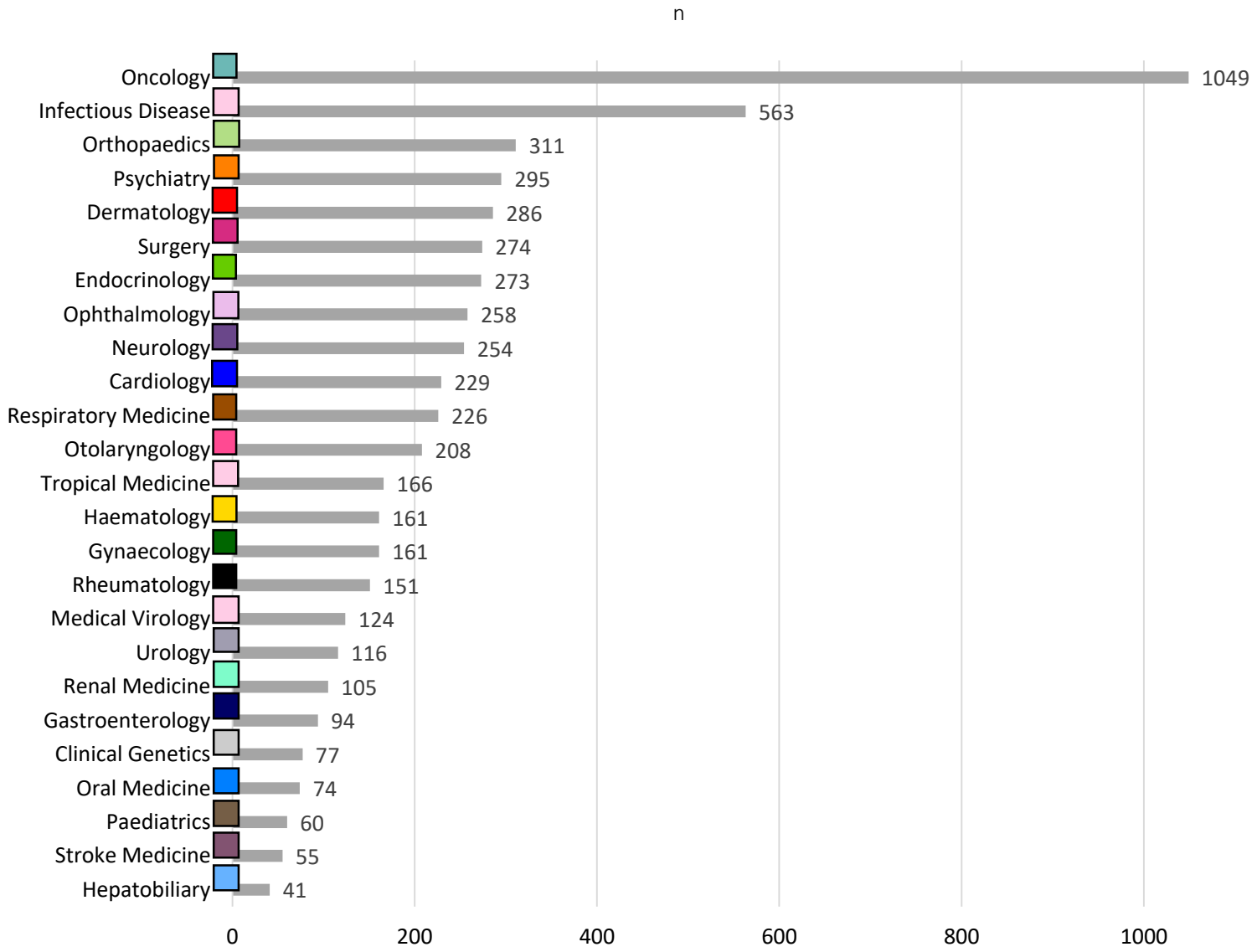
- Does not readily Classify diseases !
  - 1 disease may have many leaf codes in same chapter: clinical role in grouping e.g. Vanderbilt PheCodes
  - 1 disease may have codes ‘fragmented’ across multiple chapters.
  - 1 code can map to many diseases
  - Does not distinguish disease from non –disease codes (e.g abnormal tests)
- Does not readily engage Clinical Specialists (who create data in first place)

D-code: from ‘ICD chapter’ to clinical specialty

Chapter codes (3 characters)	Chapter
A00–B99	Certain infectious and parasitic diseases
C00–D48	Neoplasms
D50–D89	Diseases of the blood and blood-forming organs
E00–E90	Endocrine, nutritional and metabolic diseases
F00–F99	Mental and behavioural disorders
G00–G99	Diseases of the nervous system
H00–H59	Diseases of the eye and adnexa
H60–H95	Diseases of the ear and mastoid process
I00–I99	Diseases of the circulatory system
J00–J99	Diseases of the respiratory system
K00–K93	Diseases of the digestive system
L00–L99	Diseases of the skin and subcutaneous tissue
M00–M99	Diseases of the musculoskeletal system and connective tissue
N00–N99	Diseases of the genitourinary system
O00–O99	Pregnancy, childbirth and the puerperium
P00–P96	Certain conditions originating in the perinatal period
Q00–Q99	Congenital malformations, deformations and chromosomal abnormalities
R00–R99	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere
S00–T98	Injury, poisoning and certain other consequences of external causes
V01–Y98	External causes of morbidity and mortality
Z00–Z99	Factors influencing health status and contact with health services
U00–U99	Codes for special purposes



D-code: from ‘ICD chapter’ to clinical specialty



Number of unique D codes

From 12k 4 character ICD-10 codes

Using Vanderbilt PheCodes (1300 leaf codes), we are developing data informed D codes: pre-beta version.

5390 unique "Disease" type phecodes

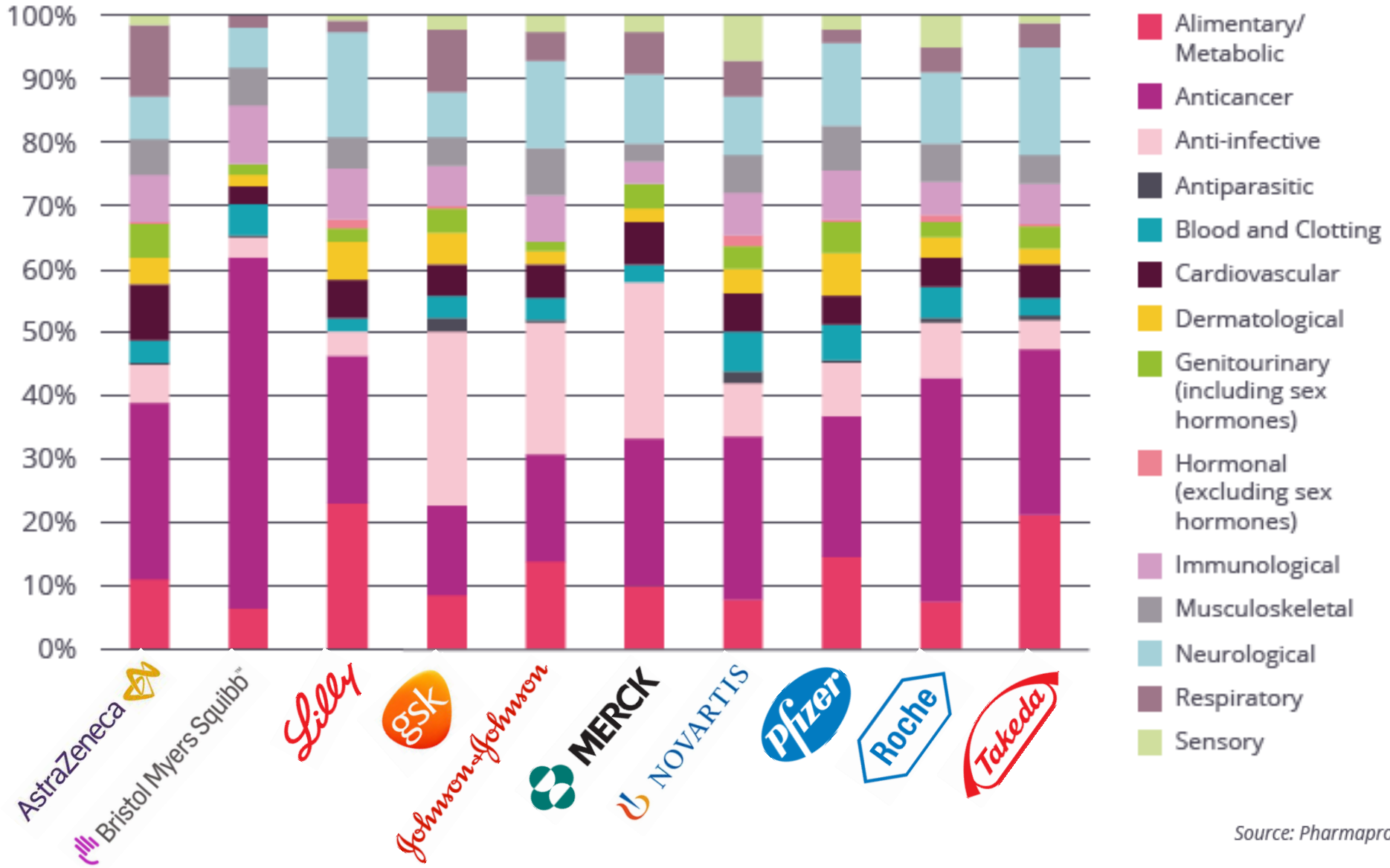
Of which
3244 are leaf codes

Of which
82 are "other specified"

Of which
43 offer specifications in rubric

Prognostic Atlas ⇋ Map to therapeutic areas of top 10 pharma companies

Pre-competitive opportunities to sponsor knowledge generation / management, across conditions in a common framework



Source: Pharmaprojects®, January 2020

D-code: clinical speciality ownership, authoring and review tools: ICD-10

- Cardiac dysrhythmias `Cardiology-Electrophysiology`
  - Paroxysmal tachycardia, unspecified (1) `Cardiology-Electrophysiology`
    - Paroxysmal supraventricular tachycardia (2) `Orphanet 2` `Cardiology-Electrophysiology`
    - Paroxysmal ventricular tachycardia (2) `Orphanet 2` `Cardiology-Electrophysiology`
  - Atrial fibrillation and flutter (4) `Orphanet 1` `Cardiology-Electrophysiology`
    - Atrial fibrillation `Cardiology-Electrophysiology`
    - Atrial flutter `Cardiology-Electrophysiology`
    - Paroxysmal atrial fibrillation (1) `Cardiology-Electrophysiology`
  - Other specified cardiac dysrhythmias (2) `Cardiology-Electrophysiology`
    - Other specified cardiac arrhythmias (1) `Orphanet 4` `Cardiology-Electrophysiology`

ICD-10 Code	ICD Label	ICD rubric	N people with Code
I49.8	Other specified cardiac arrhythmias	Brugada syndrome Long QT syndrome Rhythm disorder: coronary sinus, ectopic, nodal	59,483

OrphaCode	Orphanet Label	Map type
8022	Brugada Syndrome	Exact
10671	Torsade-de-pointes syndrome with short coupling interval	Narrow-Broad
10670	Familial short QT syndrome	Narrow-Broad
28037	GNB5-related intellectual disability-cardiac arrhythmia syndrome	Narrow-Broad

Prof Spiros Denaxas

Importance of growing use of SNOMED CT

The global language of healthcare

- ‘Common language’ across primary and secondary and tertiary care
- Semantic
- Higher clinical resolution, with >300k terms
- So may add considerable value in building a reference catalogue of disease, and prognostic atlas

…….but similar need for Clinical Speciality engagement

D-code clinical speciality ownership, authoring and review: SNOMED-CT

Dr Anoop Shah

Key to buttons for each concept

- Expand Show descendants of this concept
- Contract Hide descendants of this concept
- ? Mark as unsure (to check)
- + Add a concept
- - Remove a concept
- ++ Add a concept and all descendants
- -- Remove a concept and all descendants

Reviewing tools

Show top-level concepts only | Show all concepts

Mark all concepts as "checked" | **Mark all concepts as "unchecked"** | Show unchecked concepts only

Export to Heart failure hierarchy .csv

Expand	SNOMED CT concept	Comment	Checked	Included	
Contract	· **Heart failure with reduced ejection fraction (disorder)**	Add...		Y	? + - ++ --
	· · Heart failure with reduced ejection fraction due to cardiomyopathy (disorder)	Add...		Y	? + -
	· · Heart failure with reduced ejection fraction due to coronary artery disease (disorder)	Add...		Y	? + -
	· · Heart failure with reduced ejection fraction due to heart valve disease (disorder)	Add...		Y	? + -
	· · Heart failure with reduced ejection fraction due to myocarditis (disorder)	Add...		Y	? + -
	· High output heart failure (disorder)	Add...		Y	? + -

D-code: operationalising the ‘treatable diseasome’

- Is the disease the subject of an evidence based clinical practice guideline?

Or

- Is the disease treated with an orphan medication approved by FDA, EMA, MHRA?

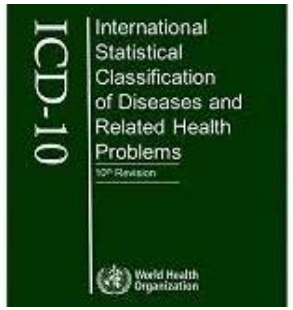
~700 distinct diseases

D-code > draft a clinically useful reference catalogue of disease (to help answer question: what do doctors do, what do hospitals do?)

Raw ICD-10	**D-Code annotation in the light of all data**
22 Chapters: heterogeneous	>> 45 clinical specialities, with review tools
12k Leaf codes used: heterogeneous	>>~3,000 disease 'leaf' codes for clinician review, *with data* for QC and engagement. If and where relative abundance makes sense, what are speciality driven priority uses?
	>> Clinical prioritisation of ~700 diseases mapping to clinical practice guideline or An approved orphan medication

Prognostic atlas: identifying and connecting underpinning elements

Diseases



Data

Doctors

GuidelineCentral

Guideline generating bodies

Charities

Dr Arturo Gonzalez-Izquierdo

Muhammad Qummer ul Arfeen

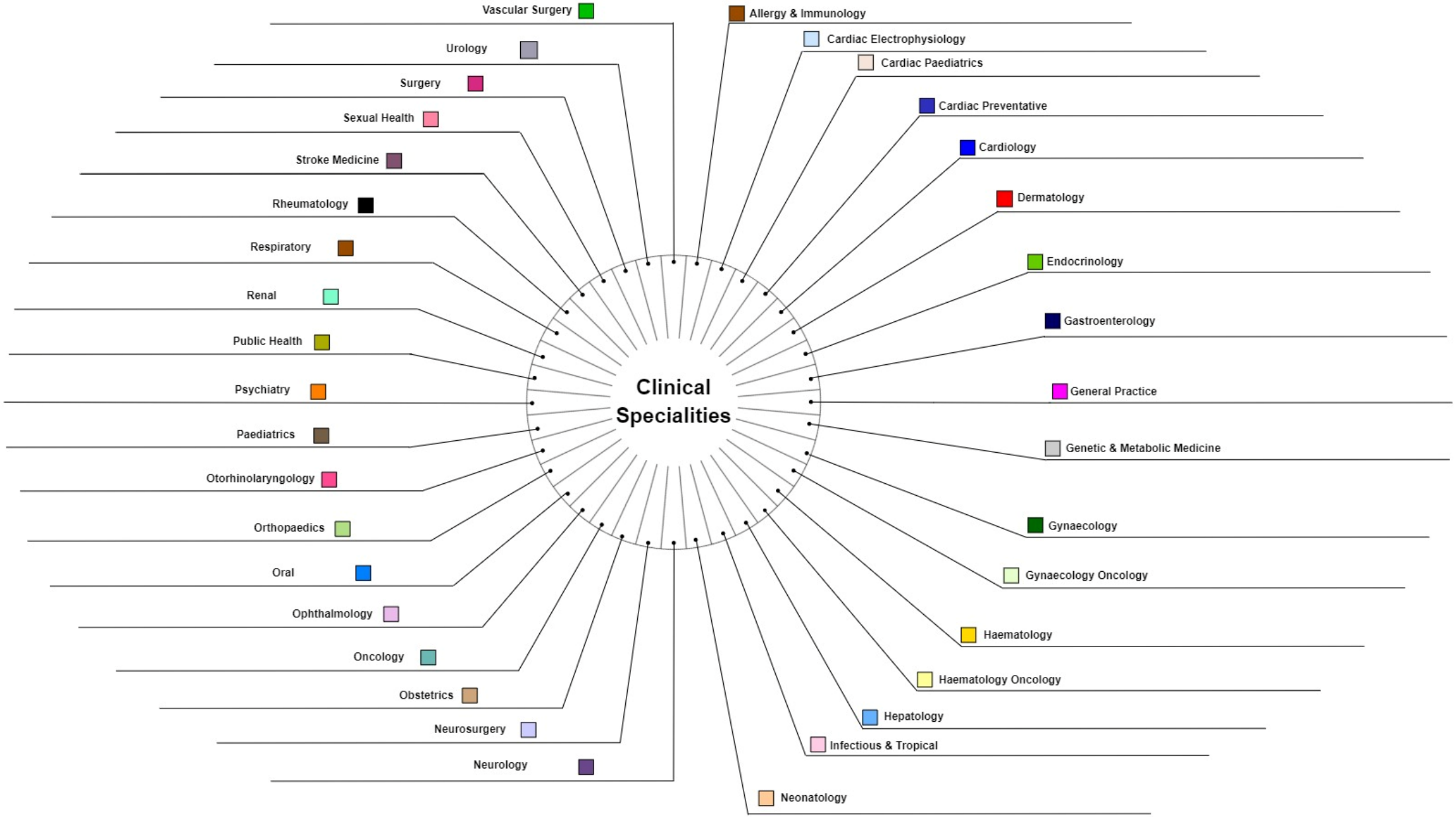
Dr Evaleen Malgapo

Natalie Fitzpatrick

Dr Serina Hayes

John Dinnewell

Izzie Harvey


Clinical Specialities
Vascular Surgery
Urology
Surgery
Sexual Health
Stroke Medicine
Rheumatology
Respiratory
Renal
Public Health
Psychiatry
Paediatrics
Otorhinolaryngology
Orthopaedics
Oral
Ophthalmology
Oncology
Obstetrics
Neurosurgery
Neurology
Allergy & Immunology
Cardiac Electrophysiology
Cardiac Paediatrics
Cardiac Preventative
Cardiology
Dermatology
Endocrinology
Gastroenterology
General Practice
Genetic & Metabolic Medicine
Gynaecology
Gynaecology Oncology
Haematology
Haematology Oncology
Hepatology
Infectious & Tropical
Neonatology

Treatable Diseasome

Vascular Surgery AAA, AVM, CarSte, MesIsc, PopAne, RAS, SAS, TOS, TxAysm, VV

Urology BPH, CalcRe, ED, Hydroc, IntCys, NmBlad, ObUrop, Peyron, Protis, RPF, Varico

Surgery AnaFis, AnoFis, BenBrea, FemHer, HaeRoi, IngHer, Phimmos, PilSin, ProAnu, UDT

Sexual Health BagVag, Chlamy, GenHer, Gonorr, HIV, HPV, LGV, MycGen, PubLic, Syphil, Tricho

Stroke Medicine CerAmy, CerAne, CerArt, HeStro, ICH, IS, LacStr, PFO, SAH, SDH, TIA

Rheumatology AnkSpo, GCA, Gout, JuvArt, OA, OP, PMR, PsorAr, RA, Raynau, ReacAr, Sjogre, SLE

Respiratory Asbest, Brecht, CysFib, DVT, IPF, Mesoth, PE, PTB, Sarcpo, SponPn

Renal AplAna, CKD, CKD, CKD, CKD2, DiaIns, IgANep, IgAVas, KidCys, LupNep, Nhrotic, PCK, PyeNep, RTA, SolKid, Wegene

Public Health Refuge, AirPol, Homles, InaHou, IVDU, LivAlo, PhyAbu, Povert, Prison, Prosti, SexAu, Unempl

Psychiatry ADHA, Alcoho, Autism, Bipola, Cannab, Cocain, Eating, IntDis, OCD, Opiods, Panic, PerDis, PTSD, Schizo

Paediatrics CerPal, CoBiAt, Galact, LacInt, LysoSt, Neurob, PilAst, PUV, RhabMy, Wilms

Otorhinolaryngology AcoNeu, BPPV, ChSinu, COM, Hearin, HyTurb, Menier, NasPol, OSA, OtoScl

Orthopaedics ColLig, CTS, LigKne, MenKne, OA-Hip, OA-Knee, Paget, ProlID, Scolio, SpoLis, SpSten

Oral Amyloi, Caries, ChPeri, CleLip, ClePal, DenImp, Kawasa, OrCand, ParCal, Peutz

Ophthalmology AMD, Behcet, Catara, Glauco, Kerato, LHON, PosUve, RetAOc, RP

Oncology CaBlad, CaBone, CaBrai, CaGall, CaKapo, Cakidn, CaLive, CaLung, CaMela, CaOeso, CaOeso, CaOeso, CaOeso, CaPanc, CaPeni, CaPros, CaTest, CaThyr

Obstetrics FeInfer, GDM, HypGrav, PIH, PlacPr, PreDV, PreEcl, RecMis, RuePsy, PlacAb

Neurosurgery AraCys, CubTun, DiaMye, MenOma,PerNer, SpiAVF, SpiAVM, SpiBif, TarTun, TriNeu

Neurology AutNeu, BellPa, CluHea, CPM, DemAlx, DemFTD, DemLew, DemVas, EpiGen, EpiMyo, EpiPet, Epiple, EsTrem, Fdysau, FSMD, GuiBar, HeAtax, HeMoSe, Huntin, IIH, INPH, IOH, MG, MitMyo, MND, MS, MSA-CBM, MSA-P, Myoclo, MyoDys, NF, Parkin, PSNP, SMA, TrMye, TubScl

Allergy & Immunology AllRhi, Anapha, Angoed, CVID, EosOes, FooAll, RhiSin, SCID, SkiAll, XLA

Cardiac Electrophysiology AF, AVB-1, AVB-2, AVB-3, LBBB, LQTS, SND, SVT, TriFas, VF-VT, WPW

Cardiac Paediatrics ASD, Coarct, ConAoR, CoPuSt, HLHS/NasPol, PDA, PDA, TetFal, TrunArt, VSD

Cardiac Preventative DysBet, FaCoHy, FmTri, FCS, FH, LPLDef, PremCo, PremMI, SCAD

Cardiology AS ACS, CAD, DCM, HCM, HF, IscCar, MI, MitSte, PPH

Dermatology AV, BCC, Eczema, LicPlan, Pegoid, Pemgus, Urtica, VenUlc, Vitili, Zoster

Endocrinology CAH, Cushin, DMMono, DMNeur, DMOpht, DMPoly, HypoTh, HypPTH, LipDys, Lipody, PAI, T1D, ThyTox

Gastroenterology AngCol, AP, Barret, ChPanc, Chron, Coelia, Divert, DuoUlc, GalSto, GasUlc, GERD

General Practice Anxiet, Asthma, COPD, Deppress, HT, HypCho, LBP, LeaDis, Obesit, T2D

Genetic & Metabolic Medicine AIP, Albini, Angelm, DownSy, Ehlers, Fabry, FragiX, Galact,GCDH, GlySto, Homocy, LysSto, Marfan, OrnCar,PKU, Pompe, TaySachs, Turner, UreaCy

Gynaecology CystOv, EndHyp, endome, Fibroi, GenHer, IncoUr, PCOS, PID, ProlPe, Vagini

Gynaecology Oncology BreaCa, CaCerv, CaOvar, CxCIS, EndoCa, VulvCa

Haematology AlThal, AplAna, AplAna, Asplen, AuHaAm, B12def, BeThal, FolAna, HaeChr, HaePhi, IDA, ITP, SCA, vonWill

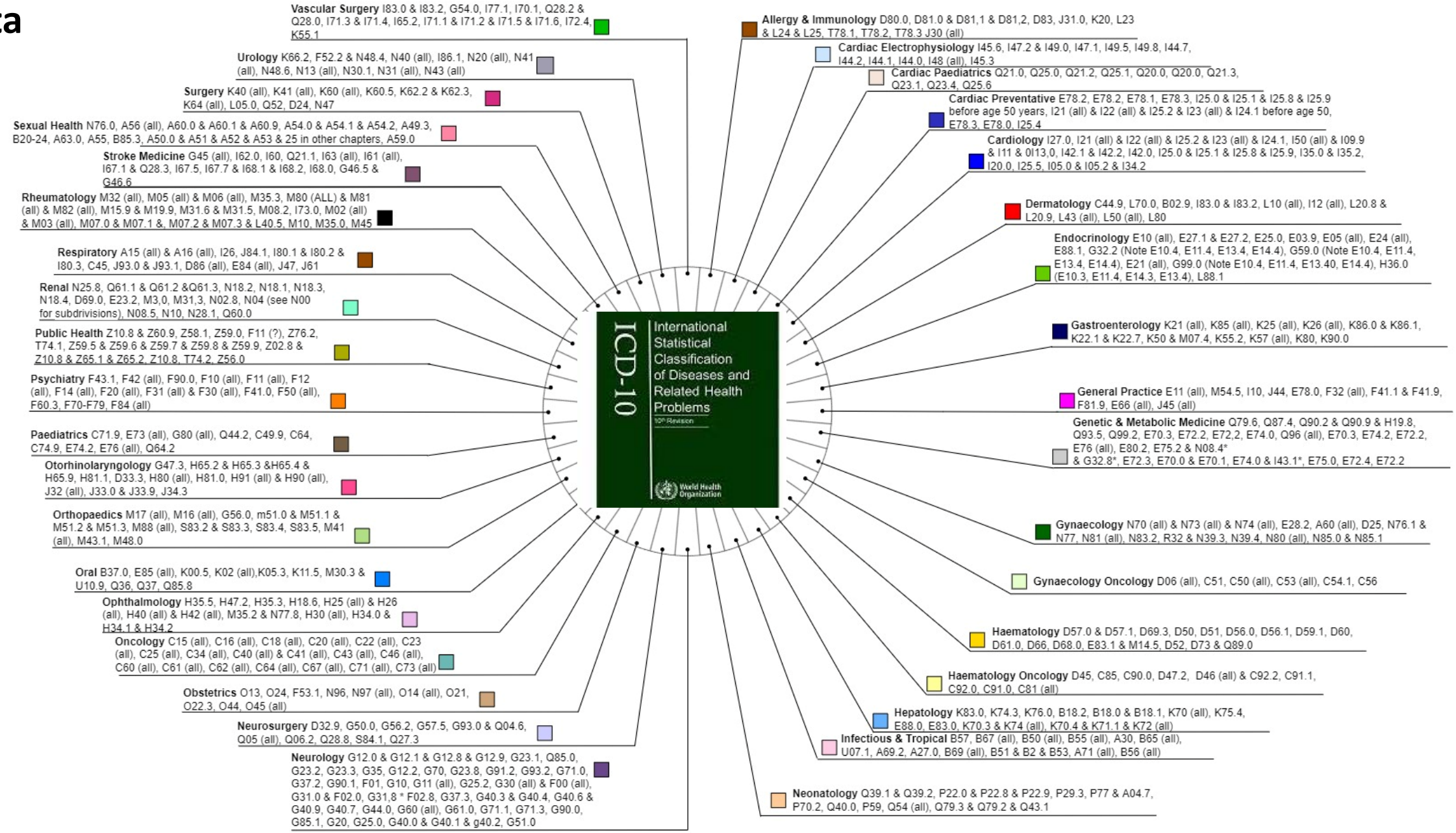
Haematology Oncology ALL, AML, CIL, CML, Hodgki, MDS, MGUS, MM, NHL, PRV

Hepatology A1TD, AIH, ALD, ChHBV, ChHCV, Cirrho, LivFai, NAFLD, PBC, PSC, Wilson

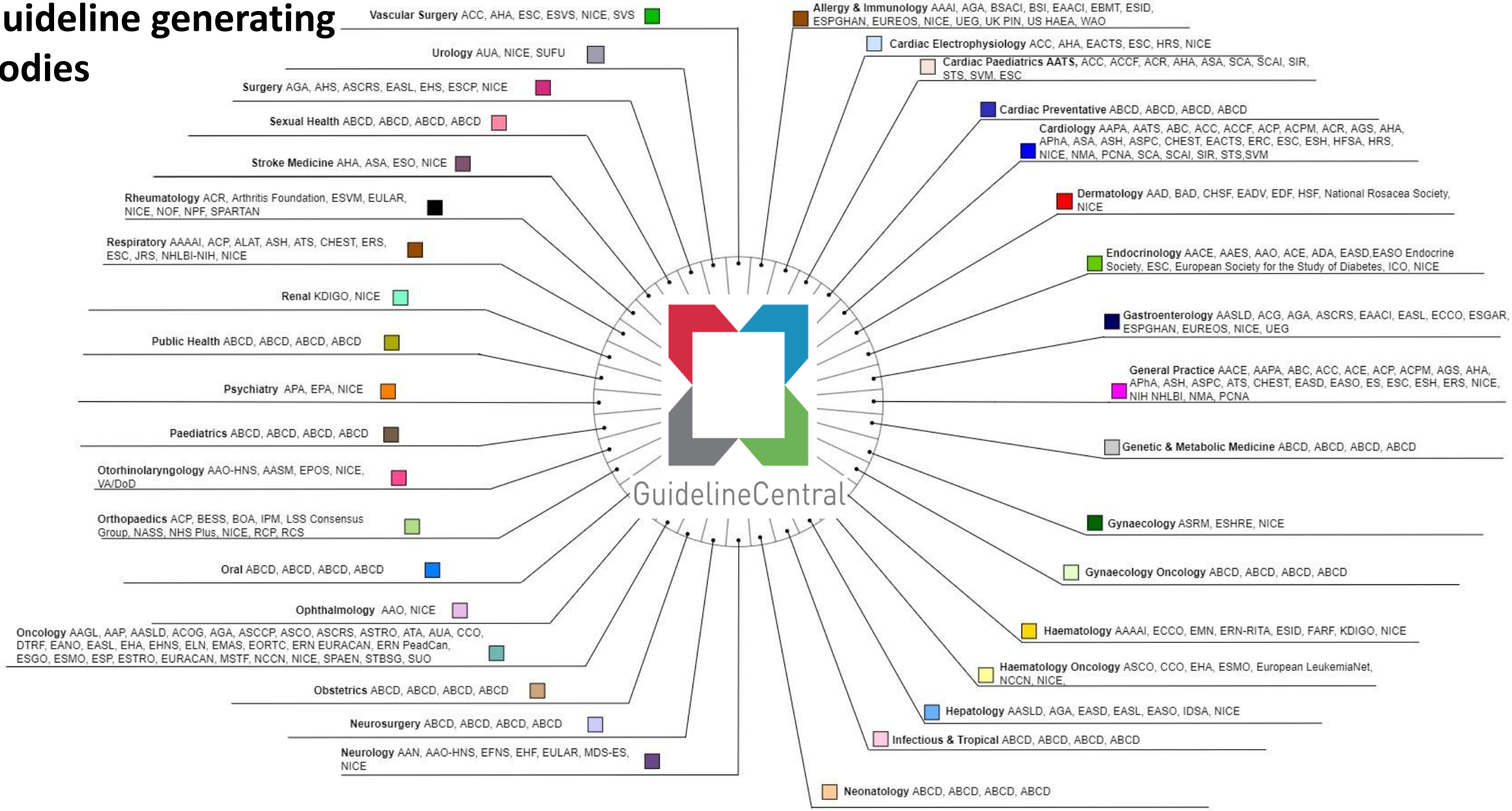
Infectious & Tropical Bilharzia, Chagas, COVID, Echino, FalMal, Leishm, Lepros, Lyme, NeuCys, OthMal, Tracho, TryAr, Weil

Neonatology CHPS, Hirsch, HySpad, IDM, NEC, neojau, Omphal, PPHT, RDS, TOF

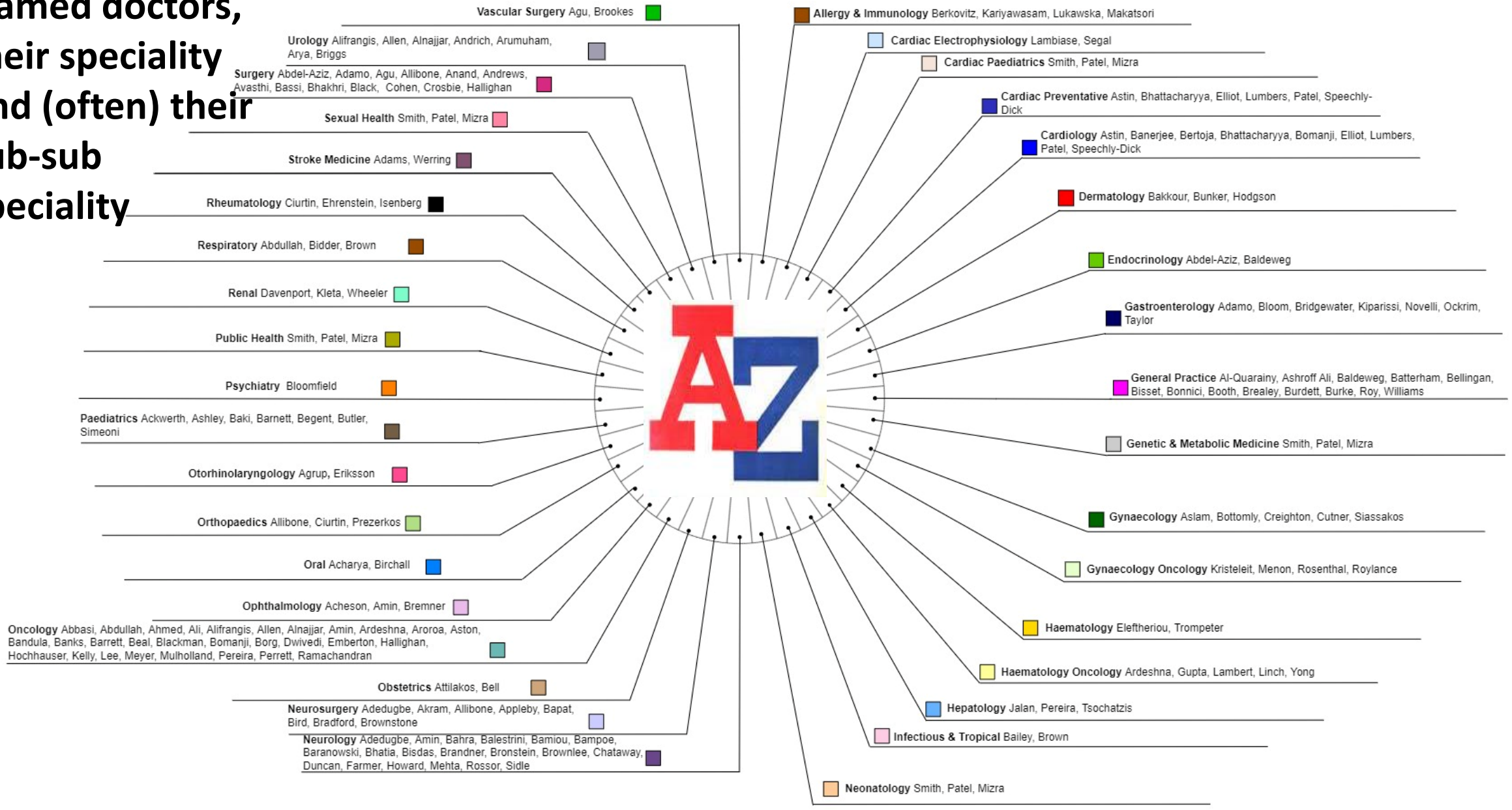
Data


ICD-10
International Statistical Classification of Diseases and Related Health Problems
10th Revision
World Health Organization
Vascular Surgery I83.0 & I83.2, G54.0, I77.1, I70.1, Q28.2 & Q28.0, I71.3 & I71.4, I65.2, I71.1 & I71.2 & I71.5 & I71.6, I72.4, K55.1
Urology K66.2, F52.2 & N48.4, N40 (all), I86.1, N20 (all), N41 (all), N48.6, N13 (all), N30.1, N31 (all), N43 (all)
Surgery K40 (all), K41 (all), K60 (all), K60.5, K62.2 & K62.3, K64 (all), L05.0, Q52, D24, N47
Sexual Health N76.0, A56 (all), A60.0 & A60.1 & A60.9, A54.0 & A54.1 & A54.2, A49.3, B20-24, A63.0, A55, B85.3, A50.0 & A51 & A52 & A53 & 25 in other chapters, A59.0
Stroke Medicine G45 (all), I62.0, I60, Q21.1, I63 (all), I61 (all), I67.1 & Q28.3, I67.5, I67.7 & I68.1 & I68.2, I68.0, G46.5 & G46.6
Rheumatology M32 (all), M05 (all) & M06 (all), M35.3, M80 (ALL) & M81 (all) & M82 (all), M15.9 & M19.9, M31.6 & M31.5, M08.2, I73.0, M02 (all) & M03 (all), M07.0 & M07.1 &, M07.2 & M07.3 & L40.5, M10, M35.0, M45
Respiratory A15 (all) & A16 (all), I26, J84.1, I80.1 & I80.2 & I80.3, C45, J93.0 & J93.1, D86 (all), E84 (all), J47, J61
Renal N25.8, Q61.1 & Q61.2 &Q61.3, N18.2, N18.1, N18.3, N18.4, D69.0, E23.2, M3,0, M31,3, N02.8, N04 (see N00 for subdrivisions), N08.5, N10, N28.1, Q60.0
Public Health Z10.8 & Z60.9, Z58.1, Z59.0, F11 (?), Z76.2, T74.1, Z59.5 & Z59.6 & Z59.7 & Z59.8 & Z59.9, Z02.8 & Z10.8 & Z65.1 & Z65.2, Z10.8, T74.2, Z56.0
Psychiatry F43.1, F42 (all), F90.0, F10 (all), F11 (all), F12 (all), F14 (all), F20 (all), F31 (all) & F30 (all), F41.0, F50 (all), F60.3, F70-F79, F84 (all)
Paediatrics C71.9, E73 (all), G80 (all), Q44.2, C49.9, C64, C74.9, E74.2, E76 (all), Q64.2
Otorhinolaryngology G47.3, H65.2 & H65.3 &H65.4 & H65.9, H81.1, D33.3, H80 (all), H81.0, H91 (all) & H90 (all), J32 (all), J33.0 & J33.9, J34.3
Orthopaedics M17 (all), M16 (all), G56.0, m51.0 & M51.1 & M51.2 & M51.3, M88 (all), S83.2 & S83.3, S83.4, S83.5, M41 (all), M43.1, M48.0
Oral B37.0, E85 (all), K00.5, K02 (all),K05.3, K11.5, M30.3 & U10.9, Q36, Q37, Q85.8
Ophthalmology H35.5, H47.2, H35.3, H18.6, H25 (all) & H26 (all), H40 (all) & H42 (all), M35.2 & N77.8, H30 (all), H34.0 & H34.1 & H34.2
Oncology C15 (all), C16 (all), C18 (all), C20 (all), C22 (all), C23 (all), C25 (all), C34 (all), C40 (all) & C41 (all), C43 (all), C46 (all), C60 (all), C61 (all), C62 (all), C64 (all), C67 (all), C71 (all), C73 (all)
Obstetrics O13, O24, F53.1, N96, N97 (all), O14 (all), O21, O22.3, O44, O45 (all)
Neurosurgery D32.9, G50.0, G56.2, G57.5, G93.0 & Q04.6, Q05 (all), Q06.2, Q28.8, S84.1, Q27.3
Neurology G12.0 & G12.1 & G12.8 & G12.9, G23.1, Q85.0, G23.2, G23.3, G35, G12.2, G70, G23.8, G91.2, G93.2, G71.0, G37.2, G90.1, F01, G10, G11 (all), G25.2, G30 (all) & F00 (all), G31.0 & F02.0, G31,8 * F02.8, G37.3, G40.3 & G40.4, G40.6 & G40.9, G40.7, G44.0, G60 (all), G61.0, G71.1, G71.3, G90.0, G85.1, G20, G25.0, G40.0 & G40.1 & g40.2, G51.0
Allergy & Immunology D80.0, D81.0 & D81,1 & D81,2, D83, J31.0, K20, L23 & L24 & L25, T78.1, T78.2, T78.3 J30 (all)
Cardiac Electrophysiology I45.6, I47.2 & I49.0, I47.1, I49.5, I49.8, I44.7, I44.2, I44.1, I44.0, I48 (all), I45.3
Cardiac Paediatrics Q21.0, Q25.0, Q21.2, Q25.1, Q20.0, Q20.0, Q21.3, Q23.1, Q23.4, Q25.6
Cardiac Preventative E78.2, E78.2, E78.1, E78.3, I25.0 & I25.1 & I25.8 & I25.9 before age 50 years, I21 (all) & I22 (all) & I25.2 & I23 (all) & I24.1 before age 50, E78.3, E78.0, I25.4
Cardiology I27.0, I21 (all) & I22 (all) & I25.2 & I23 (all) & I24.1, I50 (all) & I09.9 & I11 & 0I13.0, I42.1 & I42.2, I42.0, I25.0 & I25.1 & I25.8 & I25.9, I35.0 & I35.2, I20.0, I25.5, I05.0 & I05.2 & I34.2
Dermatology C44.9, L70.0, B02.9, I83.0 & I83.2, L10 (all), I12 (all), L20.8 & L20.9, L43 (all), L50 (all), L80
Endocrinology E10 (all), E27.1 & E27.2, E25.0, E03.9, E05 (all), E24 (all), E88.1, G32.2 (Note E10.4, E11.4, E13.4, E14.4), G59.0 (Note E10.4, E11.4, E13.4, E14.4), E21 (all), G99.0 (Note E10.4, E11.4, E13.40, E14.4), H36.0 (E10.3, E11.4, E14.3, E13.4), L88.1
Gastroenterology K21 (all), K85 (all), K25 (all), K26 (all), K86.0 & K86.1, K22.1 & K22.7, K50 & M07.4, K55.2, K57 (all), K80, K90.0
General Practice E11 (all), M54.5, I10, J44, E78.0, F32 (all), F41.1 & F41.9, F81.9, E66 (all), J45 (all)
Genetic & Metabolic Medicine Q79.6, Q87.4, Q90.2 & Q90.9 & H19.8, Q93.5, Q99.2, E70.3, E72.2, E72,2, E74.0, Q96 (all), E70.3, E74.2, E72.2, E76 (all), E80.2, E75.2 & N08.4* & G32.8*, E72.3, E70.0 & E70.1, E74.0 & I43.1*, E75.0, E72.4, E72.2
Gynaecology N70 (all) & N73 (all) & N74 (all), E28.2, A60 (all), D25, N76.1 & N77, N81 (all), N83.2, R32 & N39.3, N39.4, N80 (all), N85.0 & N85.1
Gynaecology Oncology D06 (all), C51, C50 (all), C53 (all), C54.1, C56
Haematology D57.0 & D57.1, D69.3, D50, D51, D56.0, D56.1, D59.1, D60, D61.0, D66, D68.0, E83.1 & M14.5, D52, D73 & Q89.0
Haematology Oncology D45, C85, C90.0, D47.2, D46 (all) & C92.2, C91.1, C92.0, C91.0, C81 (all)
Hepatology K83.0, K74.3, K76.0, B18.2, B18.0 & B18.1, K70 (all), K75.4, E88.0, E83.0, K70.3 & K74 (all), K70.4 & K71.1 & K72 (all)
Infectious & Tropical B57, B67 (all), B50 (all), B55 (all), A30, B65 (all), U07.1, A69.2, A27.0, B69 (all), B51 & B2 & B53, A71 (all), B56 (all)
Neonatology Q39.1 & Q39.2, P22.0 & P22.8 & P22.9, P29.3, P77 & A04.7, P70.2, Q40.0, P59, Q54 (all), Q79.3 & Q79.2 & Q43.1

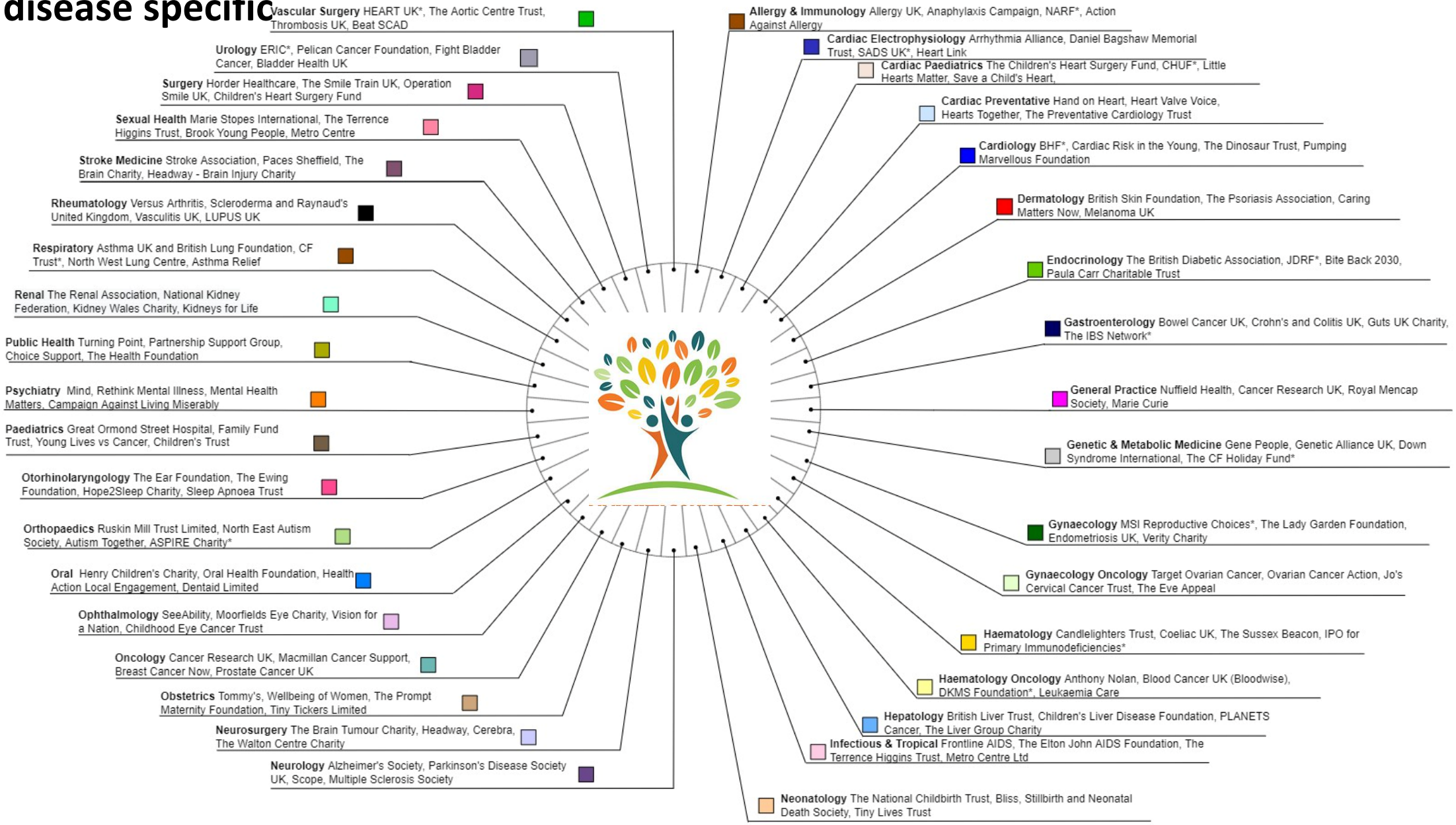
Guideline generating bodies



Named doctors, their speciality and (often) their sub-sub speciality

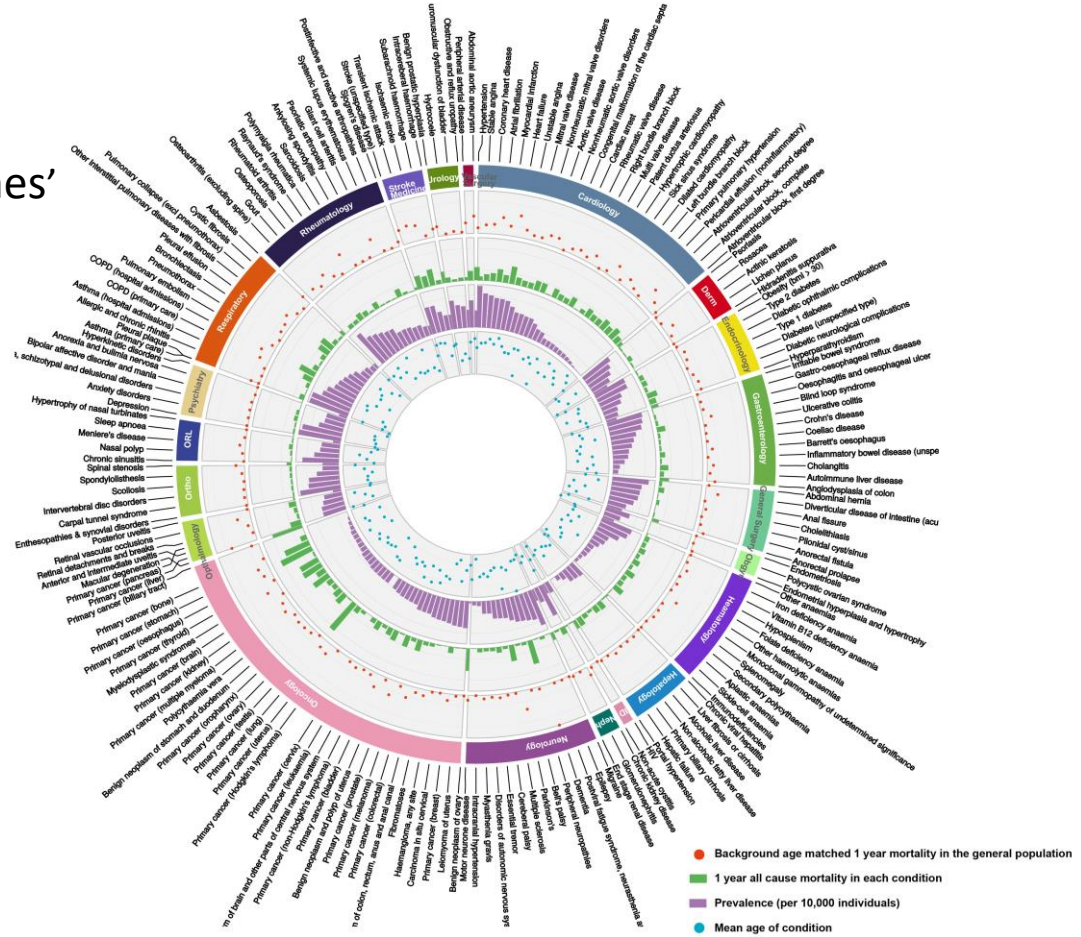


Patient organisations, charities
Often disease specific

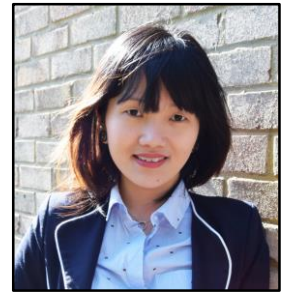


Prognostic Atlas prototype

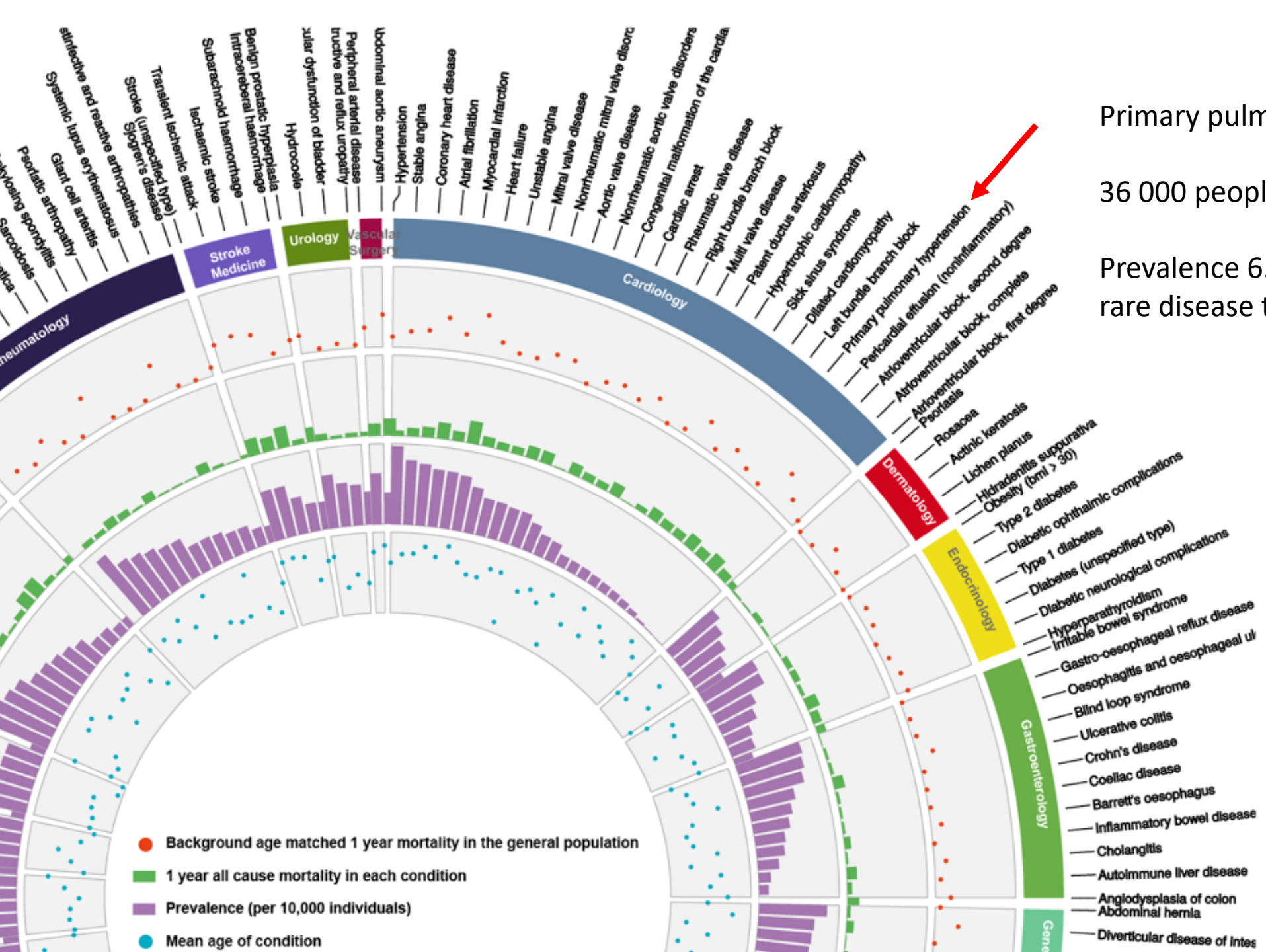
Specialities as ‘chromosomes’ (outermost track)



Small subset of conditions and specialities for display purposes



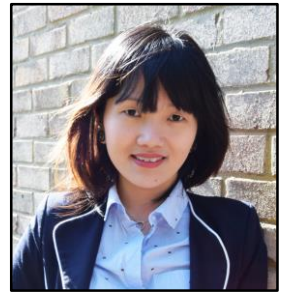
Dr Alvina Lai



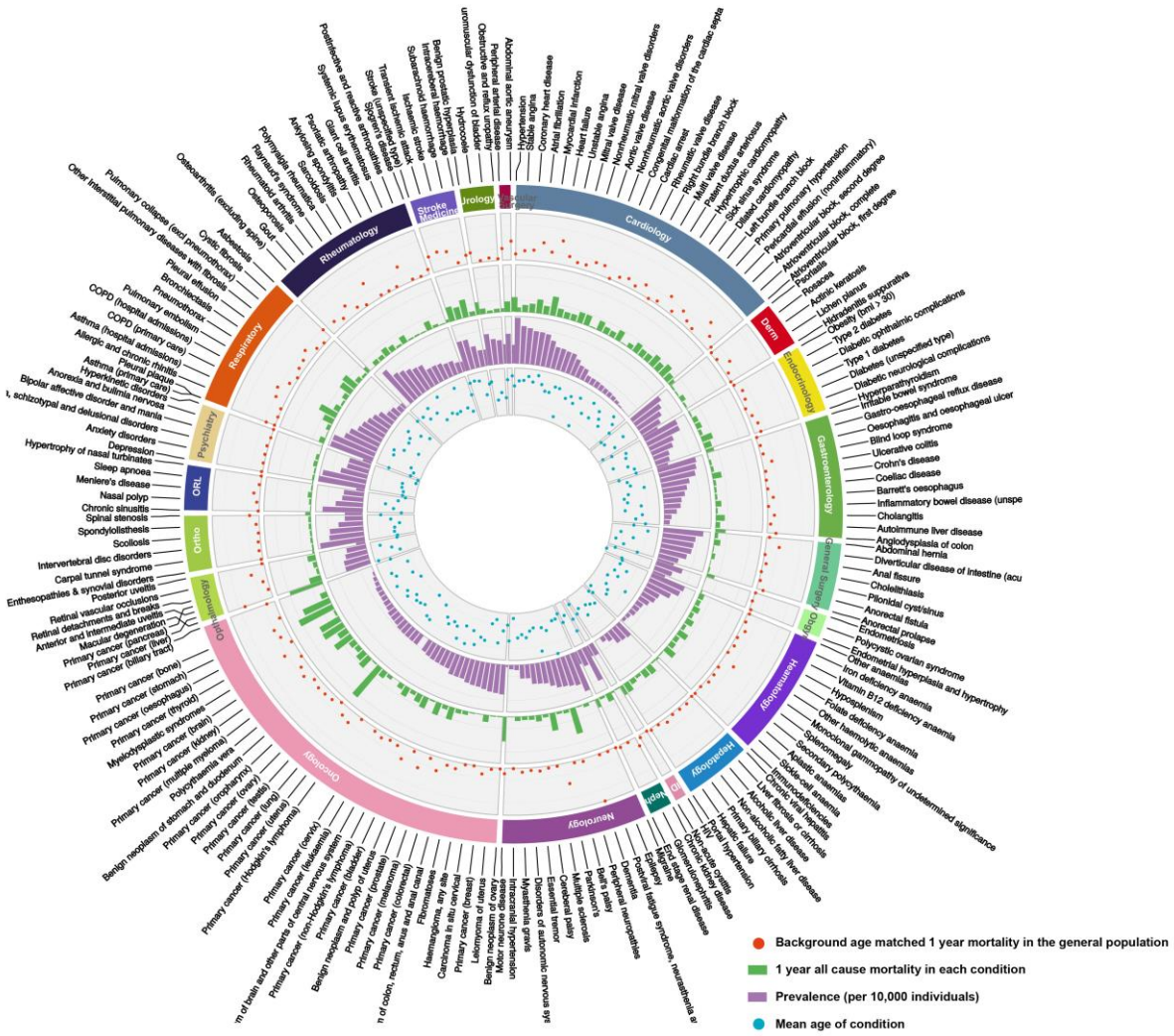
Primary pulmonary hypertension

36 000 people in England

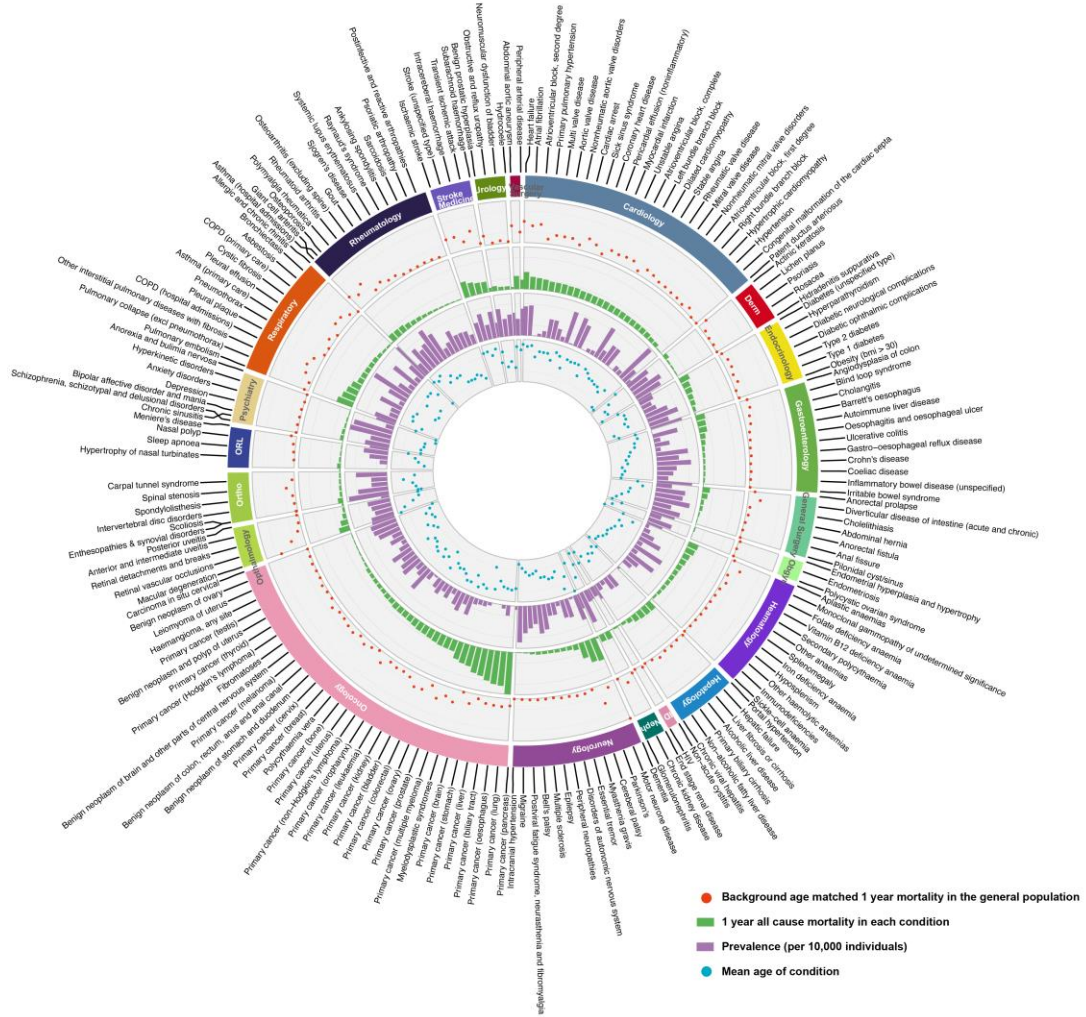
Prevalence 6.4 per 10 000 (exceeding rare disease threshold)



Dr Alvina Lai

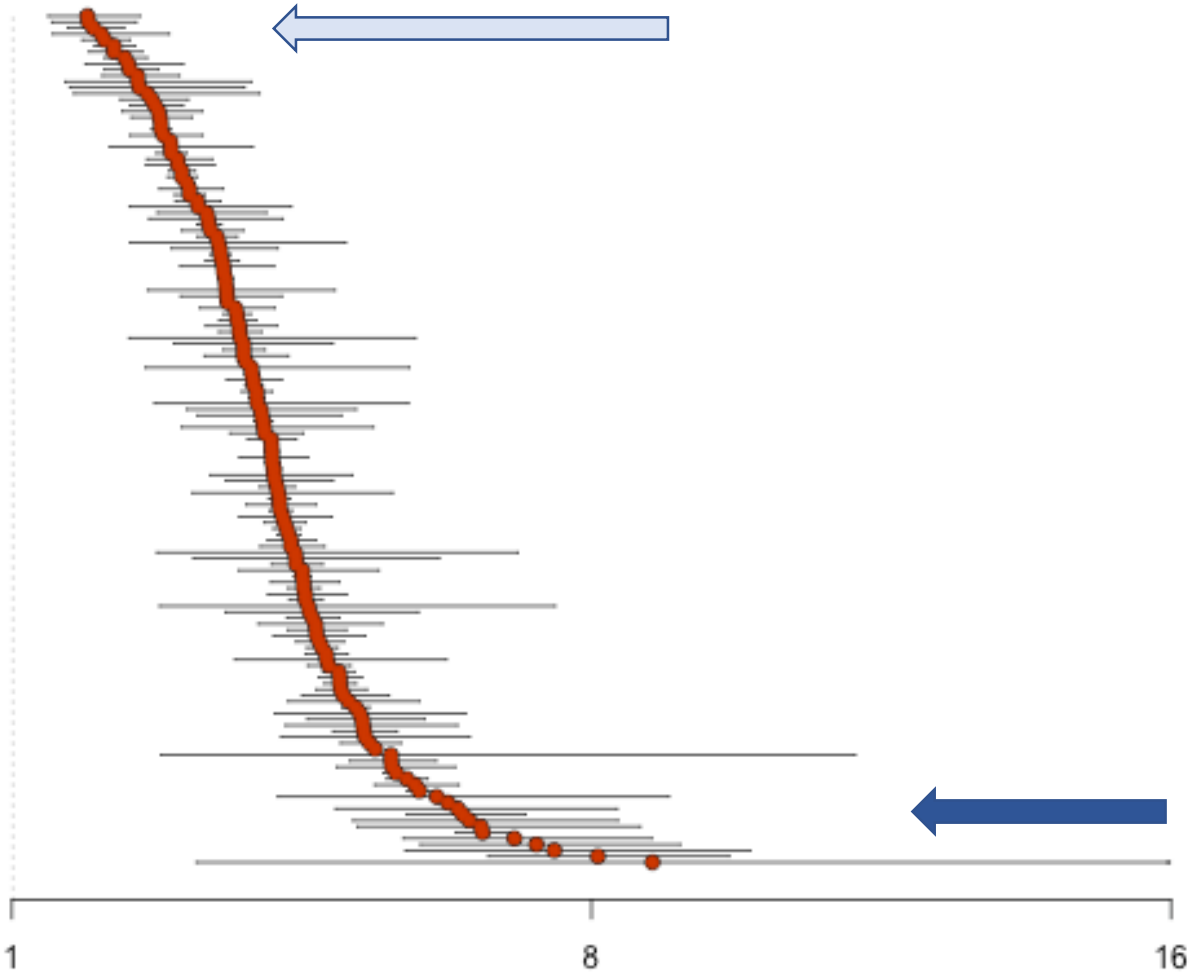


Prevalence ranked



Mortality ranked

Cross speciality Multimorbidity: Example of common cancers



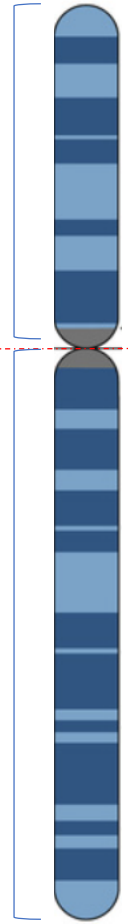
HR from 1 year any cause mortality of covid19 by phecode (95% CI)

Phecode	N deaths	N cases Phecode	N covid19	% covid19	HR (95% CI)	Crude Prevalence (10^6)
Embolism and thrombosis of renal vein	354	2533	300	11.8	1.97 (1.47-2.65)	44.74
Intracerebral haemorrhage in brain stem	441	2441	272	11.14	1.98 (1.51-2.60)	43.12
Embolism and thrombosis of vena cava	1044	6470	733	11.32	2.05 (1.72-2.45)	114.29
Polyarteritis nodosa	230	3606	367	10.1	8.01 (6.09-10.5)	63.70
Wegener's granulomatosis	569	10490	918	8.7	8.58 (7.16-10.28)	185.30
Vulval varices	21	1636	136	8.3	9.28 (3.33-25)	28.39

HR (95%CI) adjusted by Age, Sex and Ethnicity
143 phecodes with N covid cases >100 (illustration purpose)

Frequency map of disease: cardiology

	Prevalence	*N patients in NHS*	*Example diseases*
	1 in 10	5 000 000	Hypertension atrial fibrillation
‘Common’	1 in 100	500 000	Aortic stenosis Endocarditis
	1 in 1000	50 000	Primary pulmonary hypertension Hypertrophic cardiomyopathy
5 per 10 000 threshold			
	1 in 10 000	5 000	Endocardial myofibroelastosis Takayasu’s disease
‘Rare’	1 in 100 000	500	Ruptured papillary muscle Kyphoscoliotic heart disease
	1 in 1 000 000	50	Endomyocardial eosinophilia Candidal endocarditis



[

Dr Ana Torralbo

Usefulness of existing prognostic models:

from literature, one disease at a time

- Available for <5% of diseases
- Proliferate for some diseases e.g. >50 for heart failure
- Clinical practice guidelines rarely engage in prognosis, prognostic models
- In practice clinicians commonly report being ‘in the dark’
- Opportunity for engineered prognostic models across diseasome of common form
  - Near term: Age, index condition, co-existing conditions
  - Longer term: incremental prognostic value (for a given *purpose)* of molecular and other information

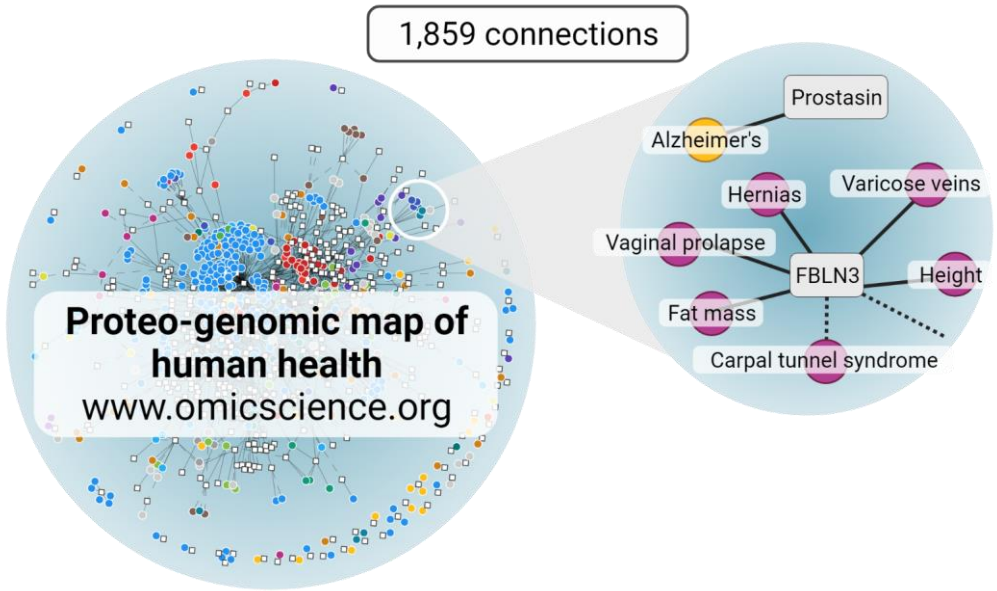
Identifying Drug Targets and Disease Mediators

RESEARCH ARTICLE SUMMARY

DISEASE GENOMICS

Mapping the proteo-genomic convergence of human diseases

Maik Pietzner†, Eleanor Wheeler†, Julia Carrasco-Zanini, Adrian Cortes, Mine Koprulu, Maria A. Wörheide, Erin Oerton, James Cook, Isobel D. Stewart, Nicola D. Kerrison, Jian'an Luan, Johannes Raffler, Matthias Arnold, Wiebke Arlt, Stephen O'Rahilly, Gabi Kastenmüller, Eric R. Gamazon, Aroon D. Hingorani, Robert A. Scott, Nicholas J. Wareham, Claudia Langenberg*



412 protein targets and

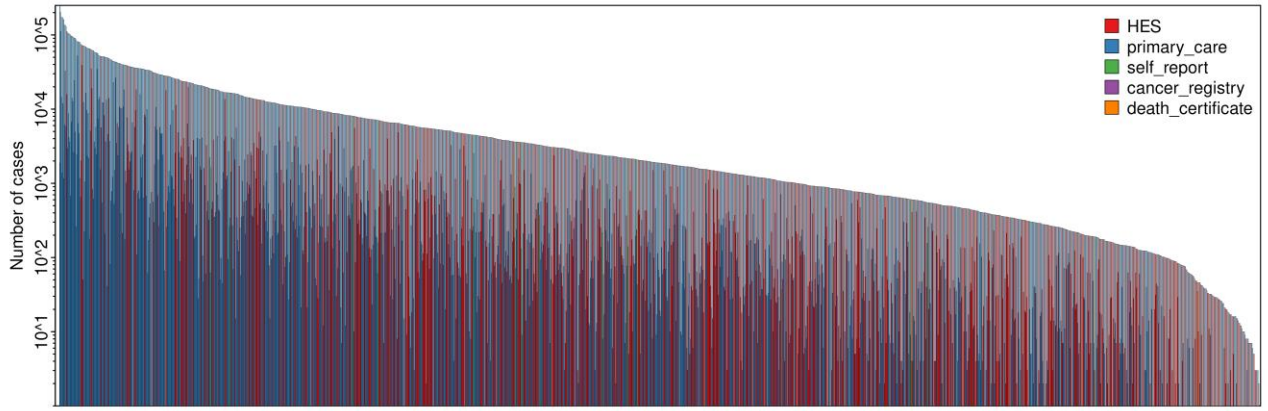
506 curated phenotypes

Harmonization across EHR resources in the UK Biobank

Dr Maik Pietzner

Dr Claudia Langenberg

N~1,500 phecodes ordered by frequency



Pietzner et al. Science 2021

Potential uses of atlas, driven by patients and specialists

- ‘How has the pandemic affected my chances of survival’
- Need for health systems to embed ‘Canaries in the mine’
- Shared decision making
- Clinical audit / quality of care / quality of data ‘Keogh principle’
- Trial feasibility and design
- ……ingenuity

Conclusion

A prognostic atlas across clinical medicine which is specialist and patient driven has become feasible in the light of current data opportunities.

If it is a duty of doctors, and a right of patients, to understand the likely course of disease, then this may be a responsibility.

Thank you