Big data indenfor sundhedsområdet og forbindelsen til ”personlig medicin”

Søren Brunak

Novo Nordisk Foundation Center for Protein Research
University of Copenhagen
soren.brunak@cpr.ku.dk

Rigshospitalet
soeren.brunak@regionh.dk

Center for Biological Sequence Analysis
Technical University of Denmark
brunak@cbs.dtu.dk


DBU

IMPRECISION MEDICINE

For every person they do help (blue), the ten highest-grossing drugs in the United States fail to improve the conditions of between 3 and 24 people (red).

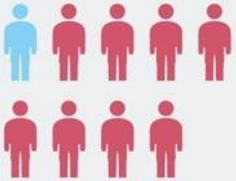
1. ABILIFY (aripiprazole)
Schizophrenia

2. NEXIUM (esomeprazole)
Heartburn

3. HUMIRA (adalimumab)
Arthritis

4. CRESTOR (rosuvastatin)
High cholesterol

5. CYMBALTA (duloxetine)
Depression



6. ADVAIR DISKUS (fluticasone propionate)
Asthma

7. ENBREL (etanercept)
Psoriasis

8. REMICADE (infliximab)
Crohn's disease

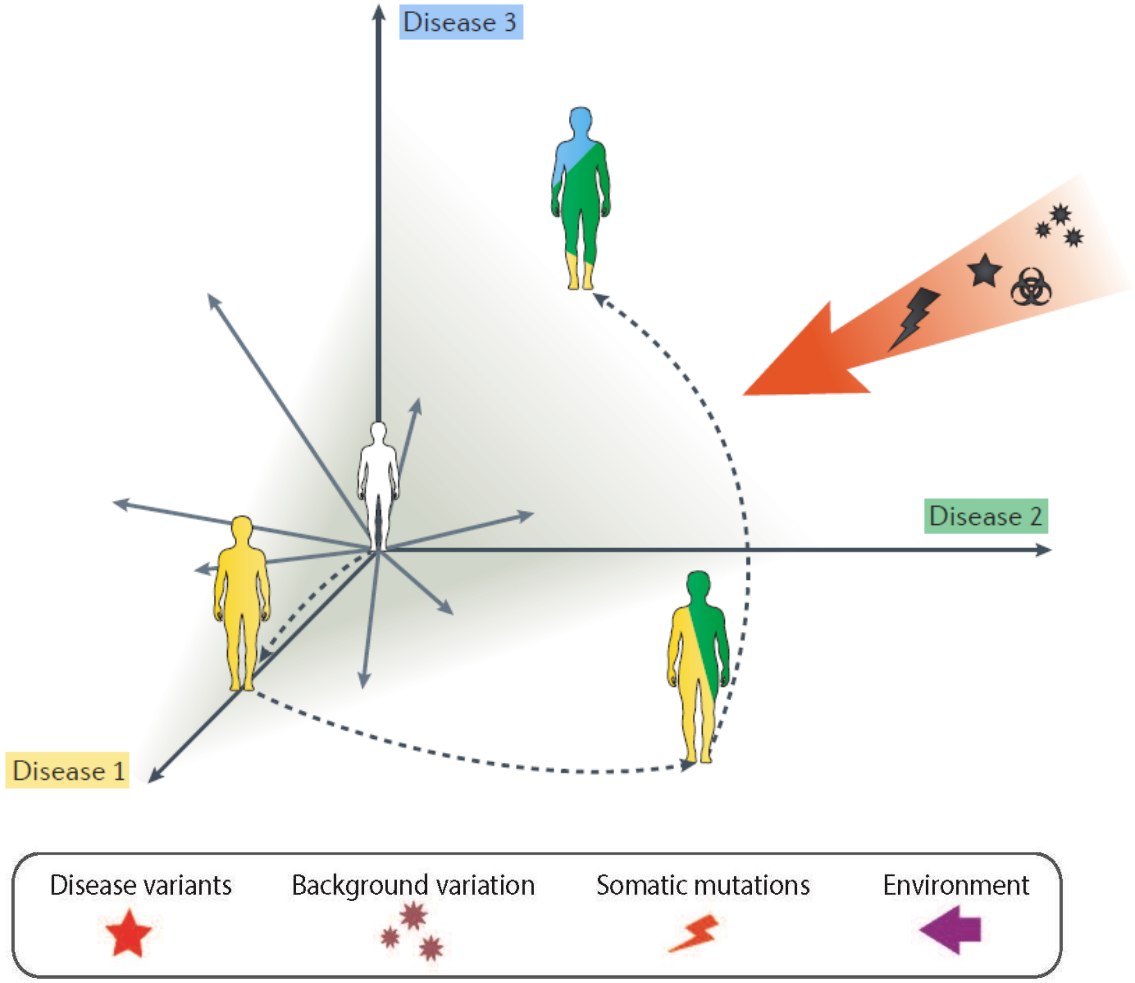
9. COPAXONE (glatiramer acetate)
Multiple sclerosis

10. NEULASTA (pegfilgrastim)
Neutropenia

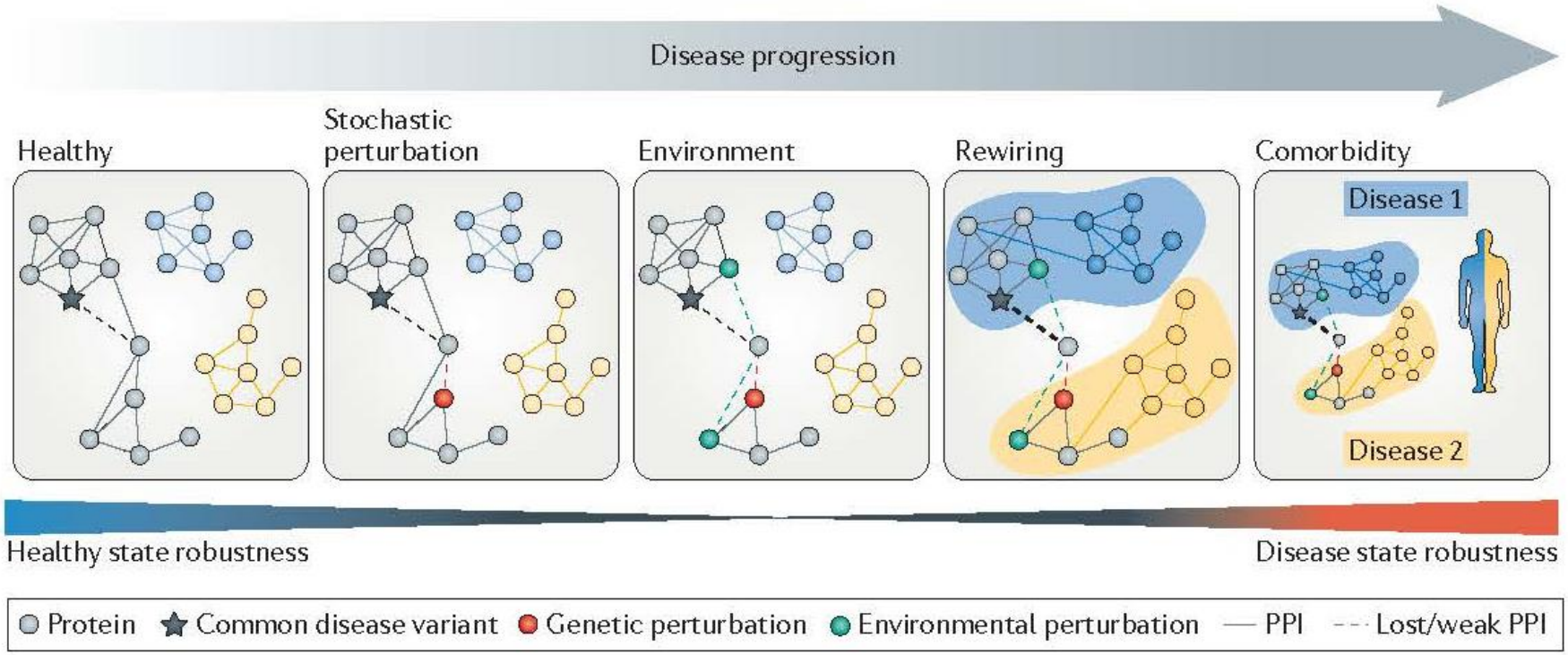
Based on published number needed to treat (NNT) figures. For a full list of references, see Supplementary Information at go.nature.com/4dr78f.

Nature 520, 609–611 (30 April 2015)

Lifelong multimorbidity journeys in disease space



Hu, Thomas & Brunak
Nature Rev. Genetics 2016



Hu, Thomas & Brunak
Nature Rev. Genetics 2016

TCCAAACCCAGGCTCTCTCCCAAACCAGTTTGCGGCAGATGGCCAGTGGAACCTCACTCTCCTCATCAGTAAAAAGGGGGCAGAGTGAGGGTCCTGAGAGCTAGTACAGGGACTGTG
TGAAGTAGACAATGCCCAGTGTTTAGCGTAAGAATCAGGGTCCAGCTGGTGCTCCCTAAACAGCAGCTGCTGTTCACTGTTGAAAGGCGCTCTGGAAGGCCAGGCGCGGTGGCTCAT
GCTTGTAATCCCAGCACTGTGGGAGGCCGAGGTGGGCGGATCACCTGAGGTAGGGAGTTCGAGACCAGCCTGACCAACGTGGAGAAACCCCATCTCTCCTAAAAATACAAAATTAGC
CAGGCGTGGTAGCACATACCTGTAATCCCAGCGACTCGGGAGGCTGAGGCAAGAGAATTGCTTGAAACCAGCAGGGGAGGTTGTGGTGAGCCAAGATCGAGCCATTGCACTCCAGCC
AGGGCAACAAGAGGCAAAATGGCGAAACTCCATCTCCGAGAAAAAAAAAAAAAAAAGAATACTTTCTGAAAGTATTTATTCATACAAATAAAGACTTGACCCATAAGGTAGGAACGCA
AATGGGCCACGGAATCACTCATTCCACAGTATACACCGAGTGCCCTTGAAGTGCTGGGCACTGCTCCAGGATTGGGGGCATATTGGTGAAAAGAGAAGCAAGCCTGCCTGCTCAGAT
GGCAGGGAATGGGGAAAAACAGGGAGACAGTTTCCTGTTTGAGATGTTGGGAGTCTGCTTCGAGTAGTATATTTACTGGAAATAGACCACTAACTTGGATGTCCCTTTTTGGAAATG
TGCCTGCGTCCAGGGCTGGGTTGGGGCCCCAATGAACTTTGGCTCTGACATAGCTGTTGCCACACTCAGTGGAACTGAATCCATGTTTGCCTTCACCCGGCATCCTTCACCCCAACT
CTCCCCGCCACAACATACATCCCATGCCAGCCTGGGGACCCTCAAAGGTGCTTCATCATTAGGTTTGTGGCTGGGTCCTACTGAAGTAAGTCTTGGCACTCAGAGGGATAGGAATTG
AATGAAGACATGAGATTCCTCTGCGGGAGGCCTCTCTAGGAAATCTGTGGACTCACACGTTTACTAATGTTGCTGCAGCCCCGCACCCACCTTGGCCTTGGGCAGCCATACTCTAGG
GCTTTTGTAACCTCTCCATGTGAGGAACTCAAATTAGACCTGGGTTTGGAGGCGGTGCTCCGAGCTGGCCTTTGGGGGAGGTTTTGTGCGAGGCATTTCCCAAGTGCTGGCAGGATT
GTGTCACAGACACAGAGTAAACTTTTGCTGGGCTCCAAGTGACCGCCCATAGTTTATTATAAAGGTGACTGCACCCTGCAGCCACCAGCACTGCCTGGCTCCACGTGCCTCCTGGTC
TCAGTATGGCGCTGTCCTGGGTTCTTACAGTCCTGAGCCTCCTACCTCTGCTGGAAGCCCAGATCCCATTGTGTGCCAACCTAGTACCGGTGCCCATCACCAACGCCACCCTGGACC
GGGTGAGTGCCTGGGCTAGCCCTGTCCTGAGCACATGGGCAGCTGCCTCCCTTCTCTGGGCTTCCCTTTACCTGCTGGCTGTGGTCGCACCCCCACTCCCAGCTCTGCCTTTTTCTC
TTCTGGGTCCCCAGGGTGAAATTCTCACCAGCCCAGGGGACTCTGGAGGCACCCCCTGCCTCCAAACACAGAAGCCTCACTGCAGAGTCCTTCACGGAGGACGGTTCTGTGCTGGGC
CTGGAGGGGCTGCCTGGGGGGCAATGACTGATCCTCAGGGTGAGCTCCTGCATGCGCACTGCCCACCAGGGGCCTCATCTCCCCATCTGCAAAATCAGGGAGAGATCTGCCTGAGTC
TCCTCCCAGCTGACAGTCAAAGATTCAGCATCAAGCCCCCATCACCAGCTCCCCCCTTCTCCCCAGATCACTGGCAAGTGGTTTTATATCGCATCGGCCTTTCGAAACGAGGAGTAC
AATAAGTCGGTTCAGGAGATCCAAGCAACCTTCTTTTACTTTACCCCCAACAAGACAGAGGACACGATCTTTCTCAGAGAGTACCAGACCCGGTGAGAGCCCCCATTCCAATGCACC
CCCGATCTCAGCTGTCTGGCCAGAAGACCTGAGCAAGTCCCTCCTTCTTCCTGGCCTTGGCCTTCCCATGGGTGGAACCGGGAGGGTTGGCTTTAATCTCCACCAGAACTCTTGCCC
CGGGACTGTGATGGGCGATTGGCCACTTCTCCTCGATAACATTACTGTTTTTCTTCCGCCTTCTGGTTGACTTTAGCCAGAACCAGTGCTTCTATAACTCCAGTTACCTGAATGTCC
AGCGGGAGAATGGGACCGTCTCCAGATACGGTGAGGGCCAGCCCTCAGGCAGGAGGGTTCACCGTGGGAACAGGGCAGGCCAGCATAAGGTGGGGGCTGGATGTAGAGCCCTGGAGG
CTTTGGGCACAGAGAAATAACCACTAACATTTTTGAGCTCTTACCACGTGCTCAGAAAAAATCCCTAAGAAGACACTGAGAGAATTAGATGAGGAAACATAAGAACAGAGACCTCAA
ATAGTTTCCCCAAGGTCACACAGCTTATAATTAGAACTAGAATTGGAACTCCAGGCTGGCTTCAGATCTGCCTCTCTCTCACGCCCTCTTTAAGATCCTTTGCAAACCAATGGTAGA
AGCCTGTATGTTGGAGAGGTGGTACCTTCAACTATGTCCCCCATCACCGCAGAGGTGGCACATGGCAGGGATCTGATGGAGCTGAACTGACATCATTTAGCATCCCGAGCCTCCTCT
CTGGGCCTCATTTTCCTCCTCTGTAAAACGGGGAGAAAGGCCCTGACAGCCACAGTCTGTGTGAGGCTCCTGAGATCTCATGTACAGAAAGTGCTTGGCGTGGAGCTGGGCACGCAG
CAGGGGCTGGGCACACGGTGGCCCAAAGGAGACCCGGGCCTTCACTGATGGGCTTTGTGGCCCCGGACACATTTCTCTTCCAGAGGGAGGCCGAGAACATGTTGCTCACCTGCGTTC
CTTAGGGACACCCCTAGGACTCCTCACCTGTAAGACAGGCACCATTGTGCCATCCCATGTTCTCACCCAGAGGCTCTTAAGACCTTGATGTTTGGTTCCTACCTGGACGATGAGAAG
AACTGGGGGCTGTCTTTCTATGGTAGGCATGCTTAGCAGCCCCAAACTCATGCCCCTCTCAGGCCTCACCCCCCATTCACCCACCCCTGGGCTGGCCCCTAGAACCCCAGCCCTCCC
TGGCCTCCGCCGGGCCCCACCATGTCCCCAGTCAGTCTCCTTGCTCCCCCTGCAGCTGACAAGCCAGAGACGACCAAGGAGCAACTGGGAGAGTTCTACGAAGCTCTCGACTGCTTG
TGCATTCCCAGGTCAGATGTCATGTACACCGACTGGAAAAAGGTAAACGCAAGGGATTGGACATTGCCCACCTTGTCCATGGCCCAACTTGGGCAGCCCCAGAGGCCCAGAGCAGGA
AAGCTGCCAGGCAAGGCTGCACAGCTAGGCAGATCTTCTGCTTTTAGGCACCTGCCTCACTGTAGGGACAGCTGAGCTCTACAGAGGCCCAGGGGTGGTGGATGAGAGCCCAGGAGG
GAGAAGTCCCTGTGAAACCAGGGAGGACCTGAAAGCTAACAGGAGGGAACAGCGTGAGCCACGGGGTTGGGGGATTGGCAATTGGAGGGGACGTAATGCGGGGAGTTACCACCTACA
GACGCGTCCCAAACCCCAGGCTTTCACCCCAACCTCCACTCCCCGCTCATTTTTAATACCCGTGCAGTGGGGAATTGATACTGTGGTTTTCAATGTCACCCACACTGCAGCACGGCC
ACAGTCACCATCCCGATTTTTGCTACAAATGAAAATTACTGTATAATGAGCTCCTTAACACTTTTCTTTAAACCTGTGTTTGGAAGACTTGTGTTGGTGTGGCCCTGTGCCCTAATA
CCTGTGAAATCACAGCACCGATGAGCTGGTTCCAATTTTTAAAATATATACATGCAGTACTTCCATGACTATTCAAAGAAAAACAATTCCTTCCATTTGCCACCTGAGATGACCACC
AGGGATGTGAACTACCTCCTGCCCCATCCCCAGCCCCAGGATCCTGGGACAGGGCTTATGAACGCAACCACTGTAGTCAGCTCACTTGATCCACAGCCTGGCACCTCCACTGTCTGG
CTAGGGAGCCTCGAATGGGTCCCAAGGCCACCCTGCTCCTCAGTTACATCATCTGCATAGTAGTGGTGGTTGTGAGGAATTCAGGAGCTGCAGCATAAGGGCCCTGCAGGTACTATG
TGCTCAGTAAATGCCAGTGGTTCTTAAGGGTCTGAGCTCCCATTGTAGAGGCAAGTAAGCTGAGGTTCAGAGAAGAAAATGACTTGCCCAAGATCACCCAGCTGGGAAGTGACAGTG
CCAGGGTTGGAGCCCTGGTTGAGCTGGTTCCACAGGCCAGAGCTCATTCTGCCCTCTCCCCGGAAGACCTCCCACCCTGTCCCCATGCCTCTGCTTCTCCCTCACCCCAATTCCCCG
CTGCCTTCTAGGATAAGTGTGAGCCACTGGAGAAGCAGCACGAGAAGGAGAGGAAACAGGAGGAGGGGGAATCCTAGCAGGACACAGCCTTGGATCAGGACAGAGACTTGGGGGCCA
TCCTGCCCCTCCAACCCGACATGTGTACCTCAGCTTTTTCCCTCACTTGCATCAATAAAGCTTCGCATCGGCCTTTCGAAACGAGGAGTACAATAAGTCGGTTCAGGAGCCCTCAGG
CAGGAGGGTTCACCGTGGGAACAGGGCAGGCCAGCATAAGGTGGGGGCTGGATGTAGAGCCCTGGAGGCTTTGGGCACAGAGGCCACCCTGGACCGGGTGAGTGCCTGGGCTAGCCC
TGTCCTGAGCACATGGGCAGCTGCCTCCCTTCTCTGGGCTTCCCTTTACCTGCTGGCTGTGGTCGCACCCCCACTCCCAGCCCCAACTCTCCCCGCCACAACATACATCCCATGCC
CAGGAGGGTTCACCGTGGGAACAGGGCAGGCCAGCATAAGGTGGGGGCTGGATGTAGAGCCCTGGAGGCTTTGGGCACAGAGGCCACCCTGGACCGGGTGAGTGCCTGGGCTAGCCC

Digital Strategy 2016-2020
Previous eGovernment Strategy 2011-2015
Simplifying regulation
Mandatory digital self-service
Mandatory Digital Post
› Strategy for Digital Welfare 2013-2020
Background
Video about digital welfare
Discussion paper
Telemedicine
Inter-ministerial Project Office
Open Data Innovation Strategy ODIS
Open Government

Strategy for Digital Welfare 2013-2020

The Danish government, Local Government Denmark and Danish Regions have jointly launched a common public sector Strategy for Digital Welfare (2013-2020).

The aim of the strategy is to accelerate the use of ICT and welfare technology in frontline public service delivery. Specifically, concrete initiatives and objectives in the strategy will speed up the use of efficient and effective digital and technological solutions in healthcare, care for the elderly, social services and education.

Also, the strategy must ensure that the public sector continually acquires new knowledge of the effects of digital technologies. And it includes a plan for testing promising technologies quickly to determine whether it would be advantageous to use them throughout Denmark.

Download the strategy here. (pdf)

New possibilities

Digital welfare means new possibilities for everyone. For the individual citizen, new digital welfare solutions can lead to better quality of life and flexibility in everyday life. Also, new technology can give citizens the possibility to actively use the resources they already have.

For instance, automated assistive devices for eating can empower persons with muscular atrophy to take more control of their own meals. And with new ICT solutions, doctors can monitor lung patients from a distance and the patients can perform their check-ups from their own homes, saving them a trip to the hospital.

For public employees, new digital and technological solutions can provide a better work environment and more efficient and effective workflows so that tasks can be solved faster and more easily. Furthermore, the demand for digital welfare solutions and the development of future welfare services has considerable potential for innovation, growth and job creation in the public sector.

Doing more with less

Read the strategy

(pdf)

Download "Digital Welfare - Empowerment, Flexibility and Efficiency (pdf)".

Press release

Read the press release "Digital welfare services will now make life easier for many citizens" from 30 September 2013.

Read the press release

Press contact

INITIATIVES UNDER THE COMMON PUBLIC-SECTOR STRATEGY FOR DIGITAL WELFARE 2013-2020

Focus area 1

DISSEMINATION OF TELEMEDICINE THROUGHOUT DENMARK

1.1 Telemedicine disseminated to patients across Denmark
- Telemedicine tested on new groups of patients
- Development of a common telemedicine infrastructure
- Dissemination of telemedicine in relevant areas
- A secure basis for telemedicine in the future

Focus area 2

EFFECTIVE COLLABORATION IN THE HEALTH AREA

2.1 Digital booking of hospital appointments

2.2 Better use of patients' own information

2.3 Use of the Shared Medication Record throughout the healthcare system

2.4 Fully digitised communication across the healthcare system

2.5 Increased service and efficiency through video interpreting and video-conferencing

Focus area 3

WELFARE TECHNOLOGY IN NURSING AND CARE

3.1 Dissemination of welfare technology in Denmark
- Assistance with lifting
- Shower toilets
- Better use of aids and appliances
- Assistive devices for eating in sheltered housing

3.2 Digital training and rehabilitation

3.3 Testing tomorrow's welfare technologies
- Smart home technologies on a large scale
- Better use of welfare technology to assist people with disabilities

Focus area 4

NEW DIGITAL PATHS IN CASE PROCESSING

4.1 Freeing up time by use of speech recognition

4.2 More knowledge about effective interventions in the social area

4.3 Municipal health tasks supported by ICT

4.4 Improved quality through data sharing

Focus area 5

DIGITAL LEARNING AND EDUCATION

5.1 Digital learning materials in primary and lower secondary schools

5.2 Digital tools in education

5.3 Digital written exams

5.4 Digital tools in daycare facilities

Focus area 6

DIGITAL COLLABORATION IN EDUCATION

6.1 Common user portal for the primary and lower secondary education

6.2 Digital education folder for exam certificates

6.3 Better sharing of digital learning materials

Focus area 7

PRECONDITIONS FOR DIGITAL WELFARE

7.1 Adequate broadband coverage for digital welfare

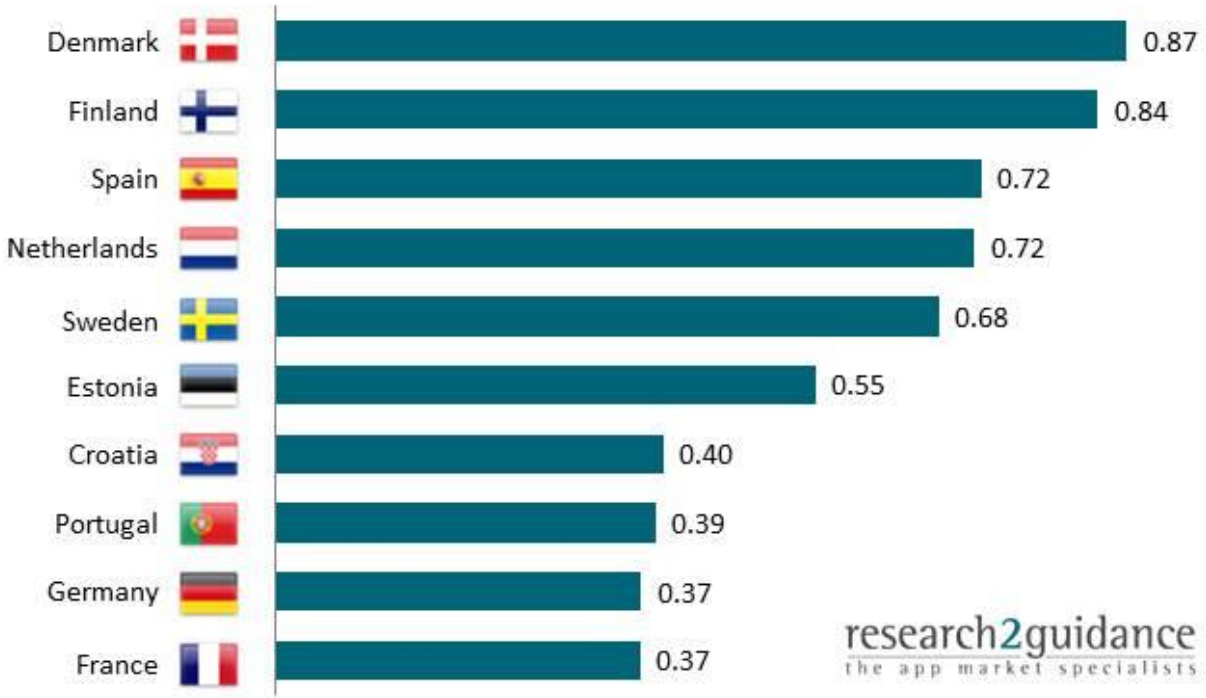
7.2 Establishing a mobile version of the digital signature for secure log-in (NemID)

7.3 Common security standards

7.4 Digital competences

DENMARK IS THE LEADING COUNTRY IN EHEALTH ADOPTION

Top 10 EU countries by eHealth adoptions of patients and doctors



eHealth adoption – doctors transferring prescription electronically, doctors electronically exchanging medical patient data with other healthcare professionals , patients making appointment via website, patients seeking online information about health

Beyond single disease analysis

(a la GWAS)

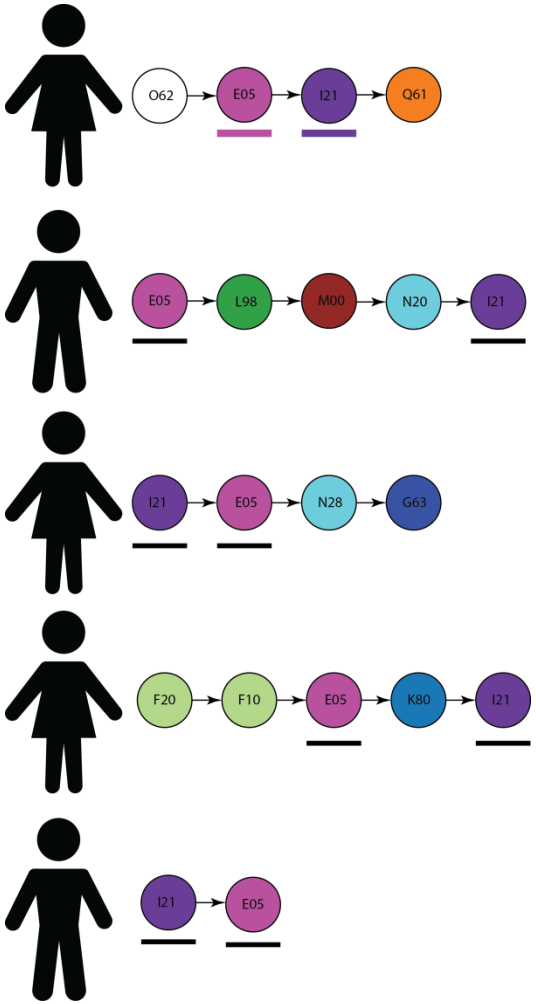
Disease-disease correlations

Disease-trajectories

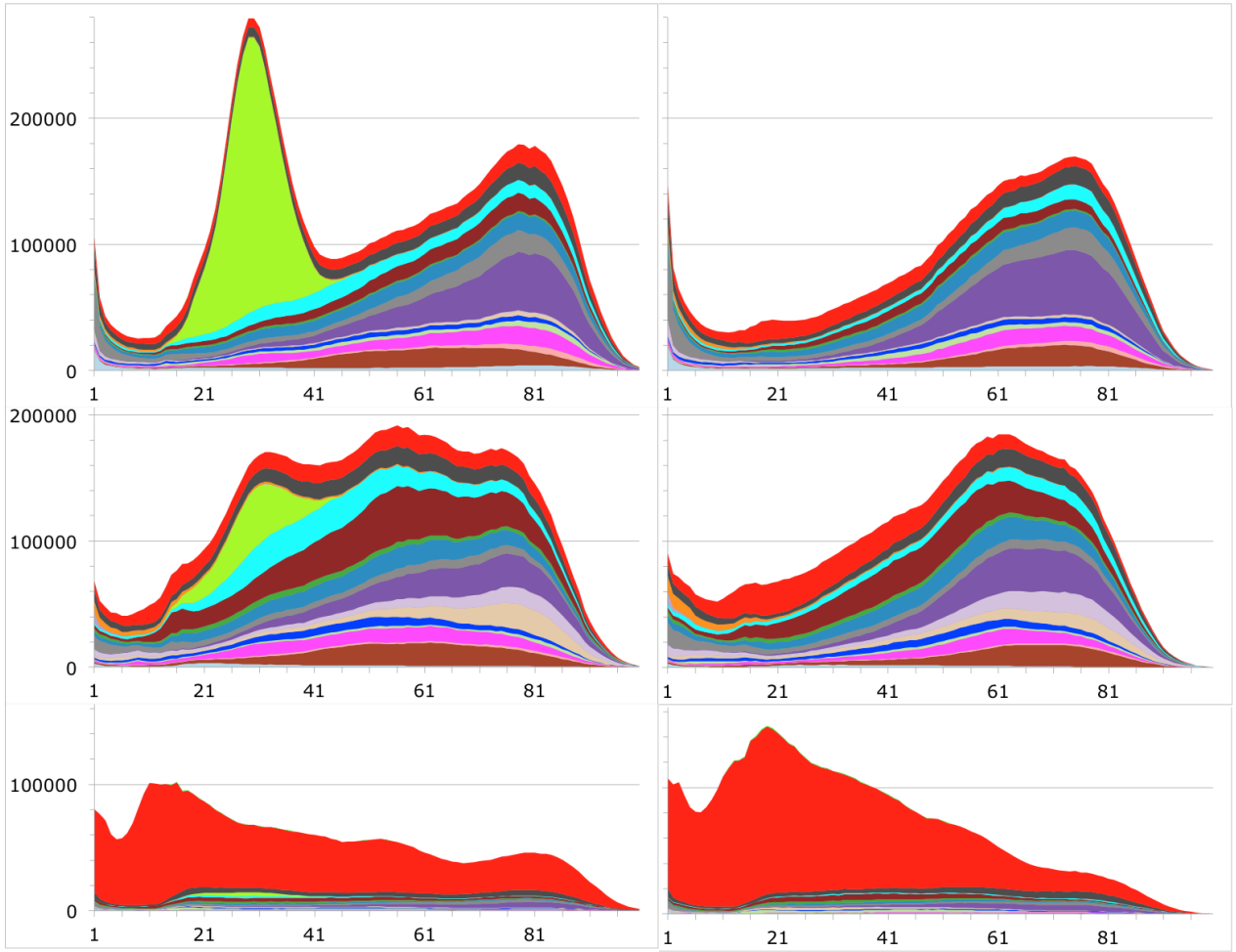
Do shared pathophysiological pathways form the basis for clinical co-occurrence, or does disease A induce disease B?

What is potentially solely genetic and what is possibly treatment related?

Diagnosis trajectories across 7 million Danish individuals (ICD-10 era, 1994-2015)



National Patient Registry (7M Danes) ICD10 diagnoses as a function of age



ICD 10 chapter coloring

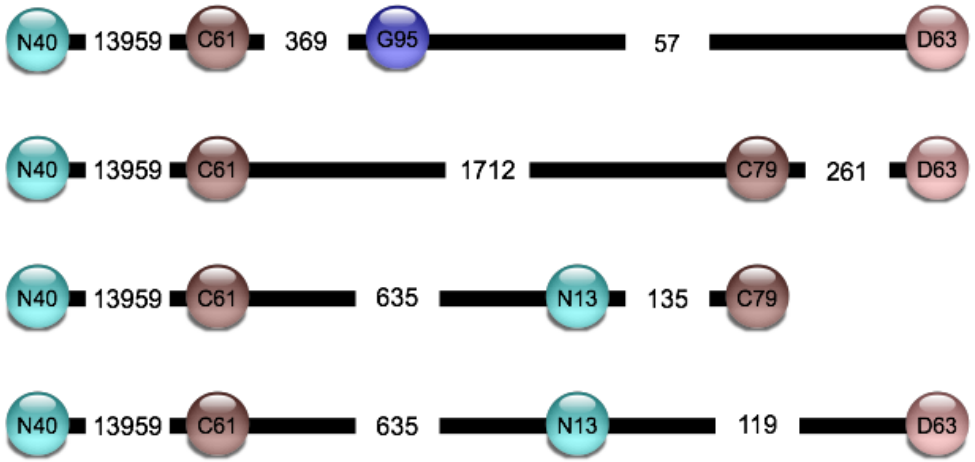
1: Certain infectious and parasitic diseases
2: Neoplasms
3: Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism
4: Endocrine, nutritional and metabolic diseases
5: Mental and behavioural disorders
6: Diseases of the nervous system
7: Diseases of the eye and adnexa
8: Diseases of the ear and mastoid process
9: Diseases of the circulatory system
10: Diseases of the respiratory system
11: Diseases of the digestive system
12: Diseases of the skin and subcutaneous tissue
13: Diseases of the musculoskeletal system and connective tissue
14: Diseases of the genitourinary system
15: Pregnancy, childbirth and the puerperium
16: Certain conditions originating in the perinatal period
17: Congenital malformations, deformations and chromosomal abnormalities
18: Symptoms, signs and abnormal clinical and aboratory findings, not elsewhere classified
19: Injury, poisoning and certain other consequences of external causes
20: External causes of morbidity and mortality

AB Jensen et al., Nature Comm., 2014

6-7 million trajectories

?

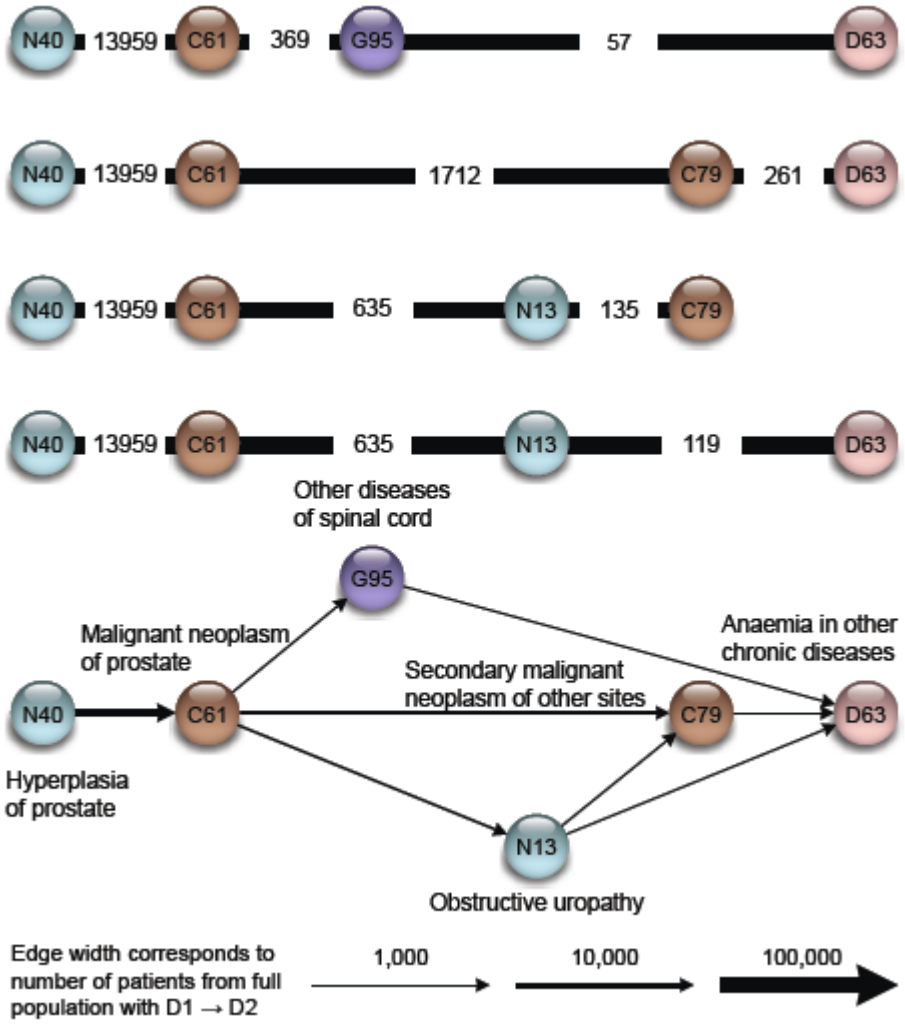
6.2 million individual trajectories condensed into 1,171 “recurrent” ones



• • •

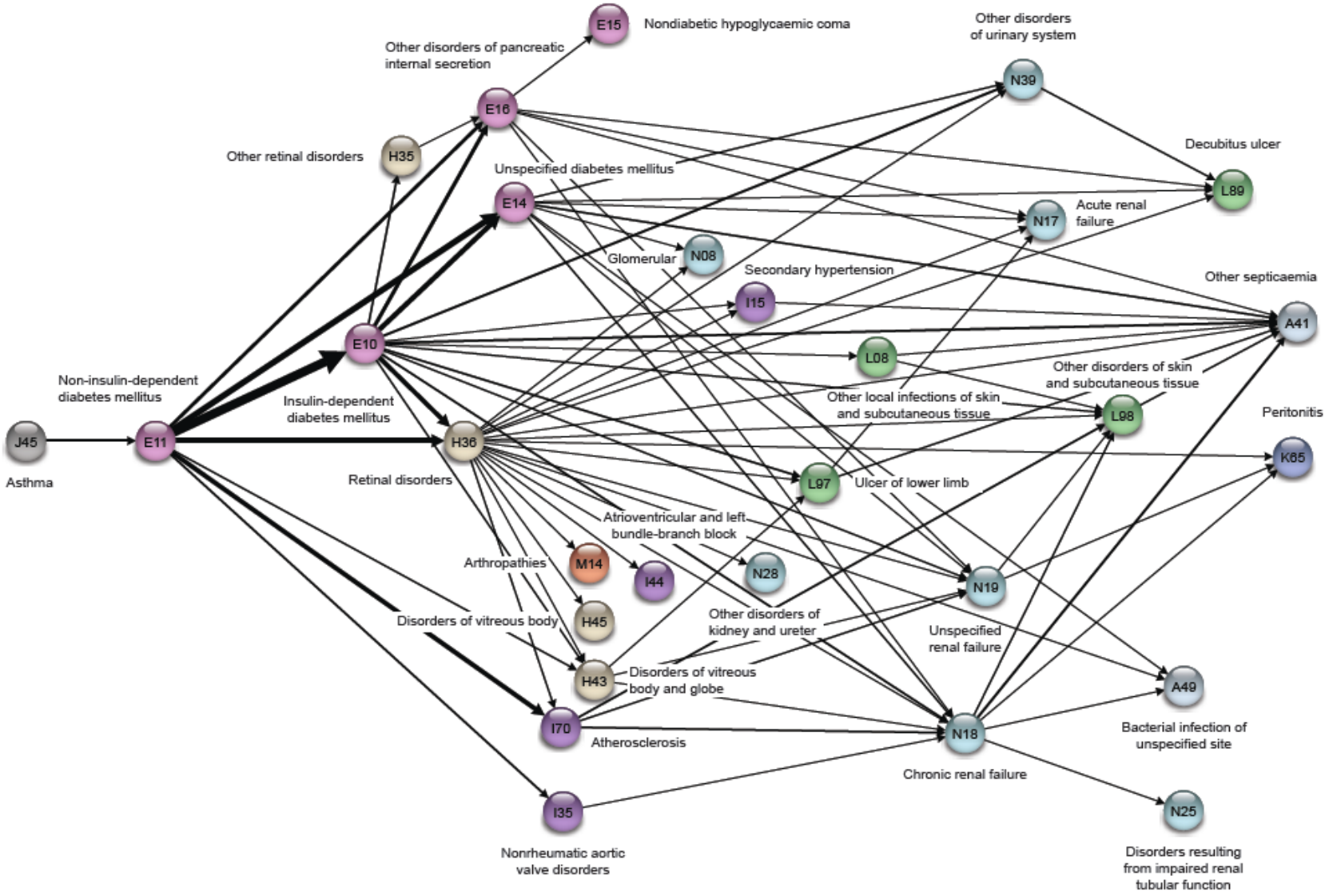
AB Jensen et al.,
Nature Comm., 2014

Disease trajectories and trajectory-cluster for prostate cancer



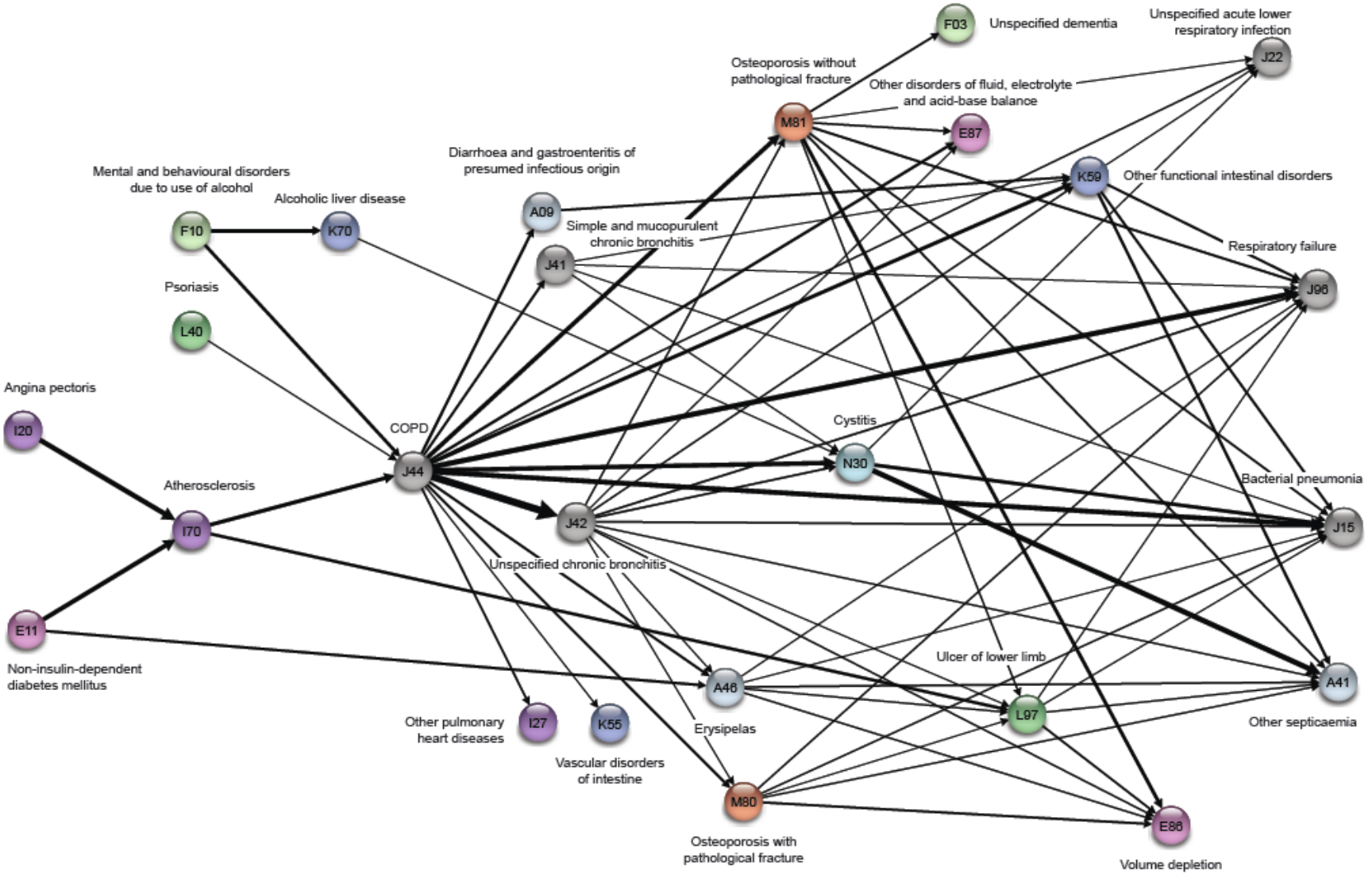
AB Jensen et al., Nature Comm., 2014

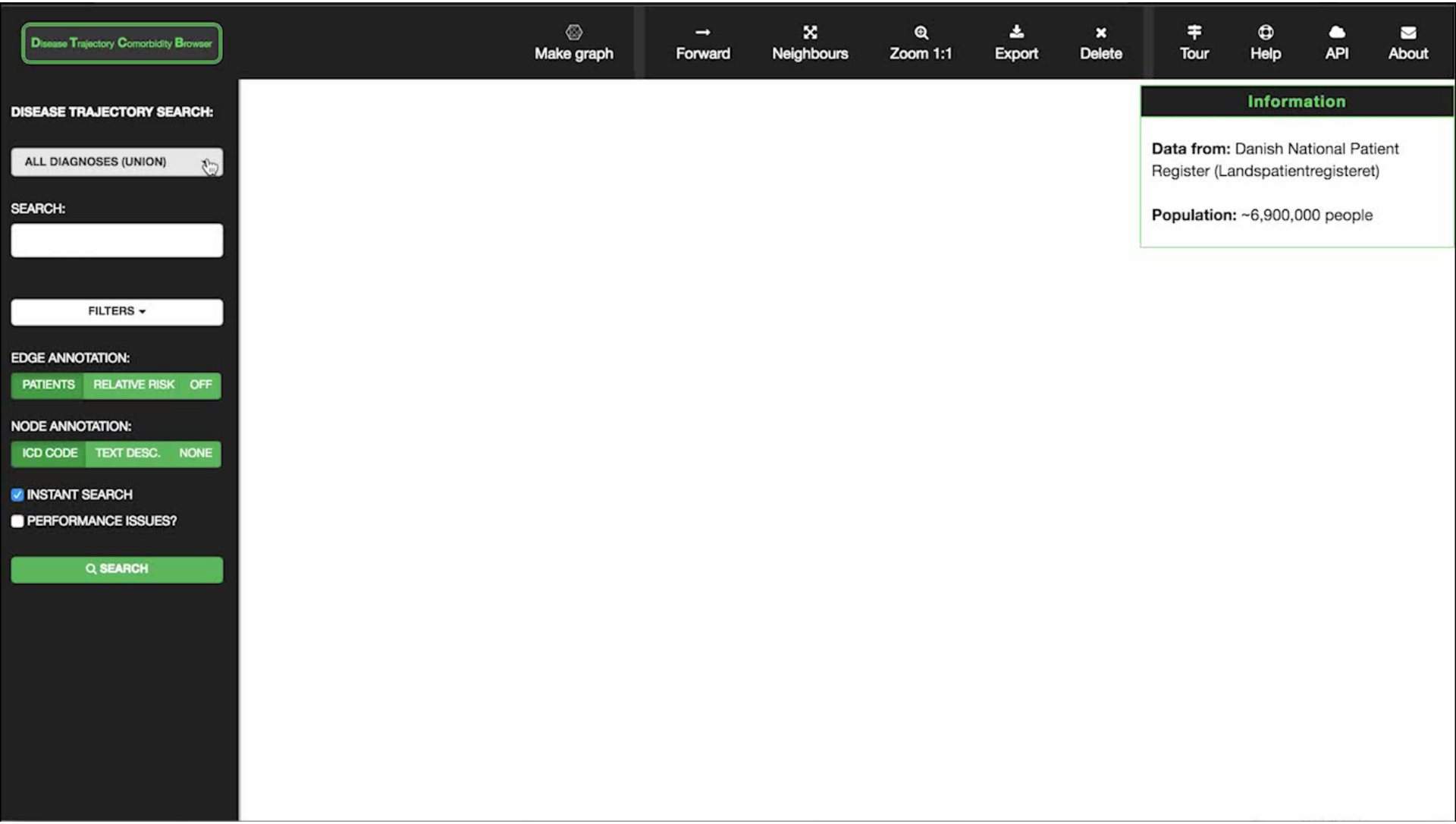
Diabetes trajectory network



COPD trajectory cluster

with five preceding diagnoses leading to COPD and some of the possible outcomes




Disease Trajectory Comorbidity Browser
Make graph
Forward
Neighbours
Zoom 1:1
Export
Delete
Tour
Help
API
About
DISEASE TRAJECTORY SEARCH:
ALL DIAGNOSES (UNION)
SEARCH:
FILTERS
EDGE ANNOTATION:
PATIENTS
RELATIVE RISK
OFF
NODE ANNOTATION:
ICD CODE
TEXT DESC.
NONE
INSTANT SEARCH
PERFORMANCE ISSUES?
SEARCH
Information
Data from: Danish National Patient Register (Landspatientregisteret)
Population: ~6,900,000 people

Nature Genetics, 48:510-518, 2016

nature genetics

Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci

David Ellinghaus[1,49], Luke Jostins[2], Sarah L Spain[2], Adrian Cortes[3,4], Jörn Bethune[1], Buhm Han[5], Yu Rang Park[6], Soumya Raychaudhuri[7–10], Jennie G Pouget[11,12], Matthias Hübenthal[1], Trine Folseraas[13–16], Yunpeng Wang[17], Tonu Esko[18–20], Andres Metspalu[18], Harm-Jan Westra[7–10], Lude Franke[21], Tune H Pers[7,20,22,23], Rinse K Weersma[24], Valerie Collij[24], Mauro D'Amato[25,26], Jonas Halfvarson[27], Anders Boeck Jensen[28], Wolfgang Lieb[29,30], Franziska Degenhardt[31,32], Andreas J Forstner[31,32], Andrea Hofmann[31,32], The International IBD Genetics Consortium (IIBDGC)[33], International Genetics of Ankylosing Spondylitis Consortium (IGAS)[33], International PSC Study Group (IPSCSG)[33], Genetic Analysis of Psoriasis Consortium (GAPC)[33], Psoriasis Association Genetics Extension (PAGE)[33], Stefan Schreiber[1,34], Ulrich Mrowietz[35], Brian D Juran[36], Konstantinos N Lazaridis[36], Søren Brunak[28], Anders M Dale[17,37], Richard C Trembath[38], Stephan Weidinger[35], Michael Weichenthal[35], Eva Ellinghaus[1], James T Elder[39,40], Jonathan N W N Barker[41], Ole A Andreassen[42,43], Dermot P McGovern[44,45], Tom H Karlsen[13–16], Jeffrey C Barrett[2], Miles Parkes[46], Matthew A Brown[47,48,50] & Andre Franke[1,50]

We simultaneously investigated the genetic landscape of ankylosing spondylitis, Crohn's disease, psoriasis, primary sclerosing cholangitis and ulcerative colitis to investigate pleiotropy and the relationship between these clinically related diseases. Using high-density genotype data from more than 86,000 individuals of European ancestry, we identified 244 independent multidisease signals, including 27 new genome-wide significant susceptibility loci and 3 unreported shared risk loci. Complex pleiotropy was supported when contrasting multidisease signals with expression data sets from human, rat and mouse together with epigenetic and expressed enhancer profiles. The comorbidities among the five immune diseases were best explained by biological pleiotropy rather than heterogeneity (a subgroup of cases genetically identical to those with another disease, possibly owing to diagnostic misclassification, molecular subtypes or excessive comorbidity). In particular, the strong comorbidity between primary sclerosing cholangitis and inflammatory bowel disease is likely the result of a unique disease, which is genetically distinct from classical inflammatory bowel disease phenotypes.

Genome-wide association studies (GWAS) have shown overlap in the genetic susceptibility to human diseases that affect a range of

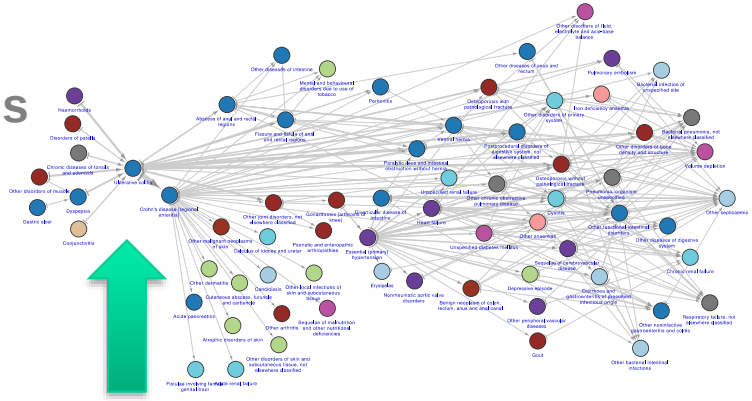
In this study, we combined Immunochip genotype data for 52,262 cases and 34,213 controls of European ancestry, currently the largest available

Immunochip genotype data: 52,262 cases

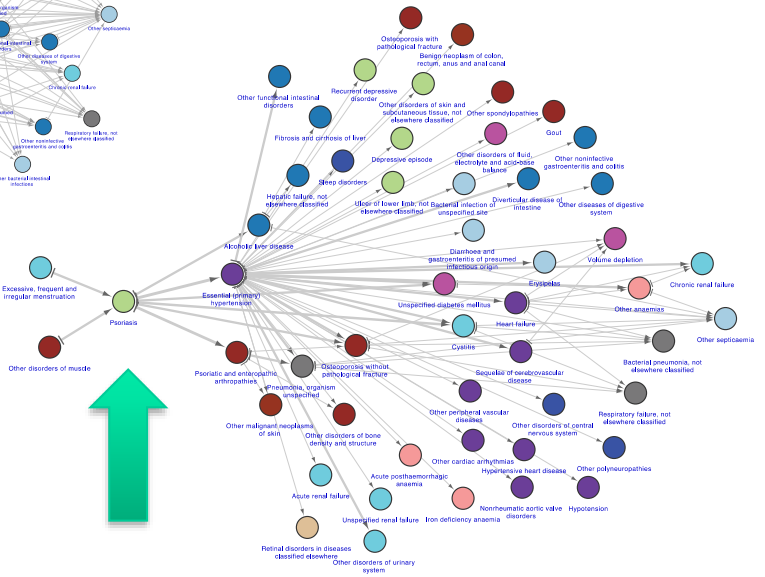
Ankylosing spondylitis (8,726)
Crohn's disease (19,085)
Psoriasis (6,530)
Primary sclerosing cholangitis (3,408)
Ulcerative colitis (14,413)

34,213 healthy controls

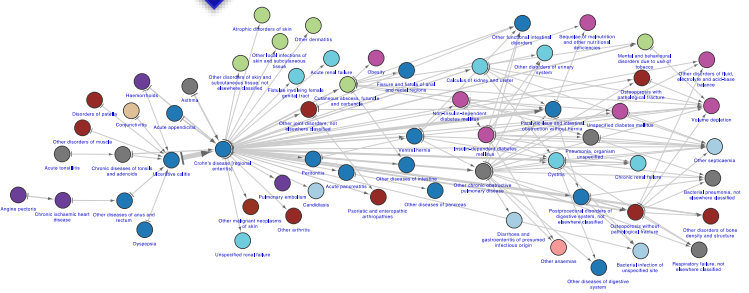
Ulcerative colitis



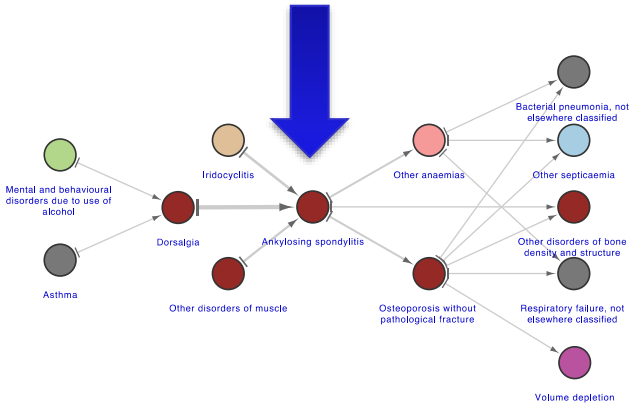
Psoriasis



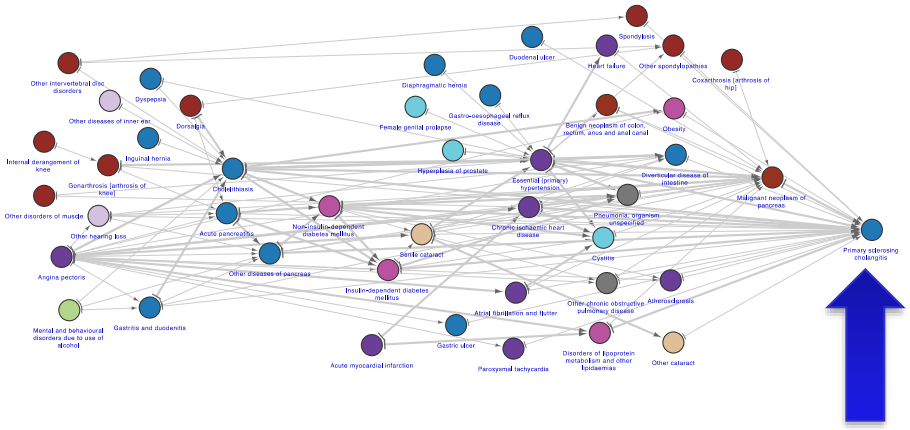
Crohn's disease



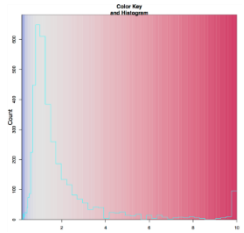
Ankylosing spondylitis



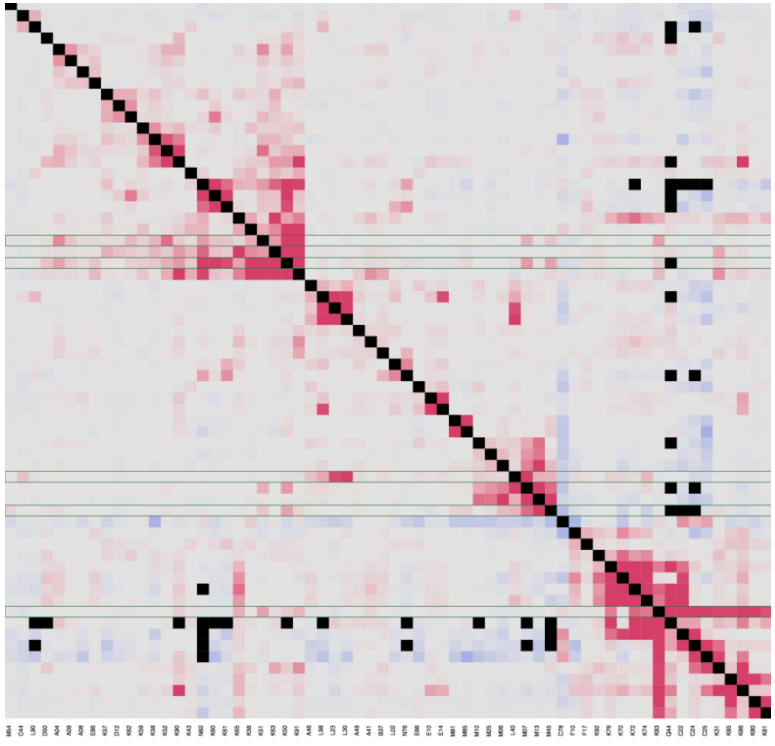
Sclerosing cholangitis



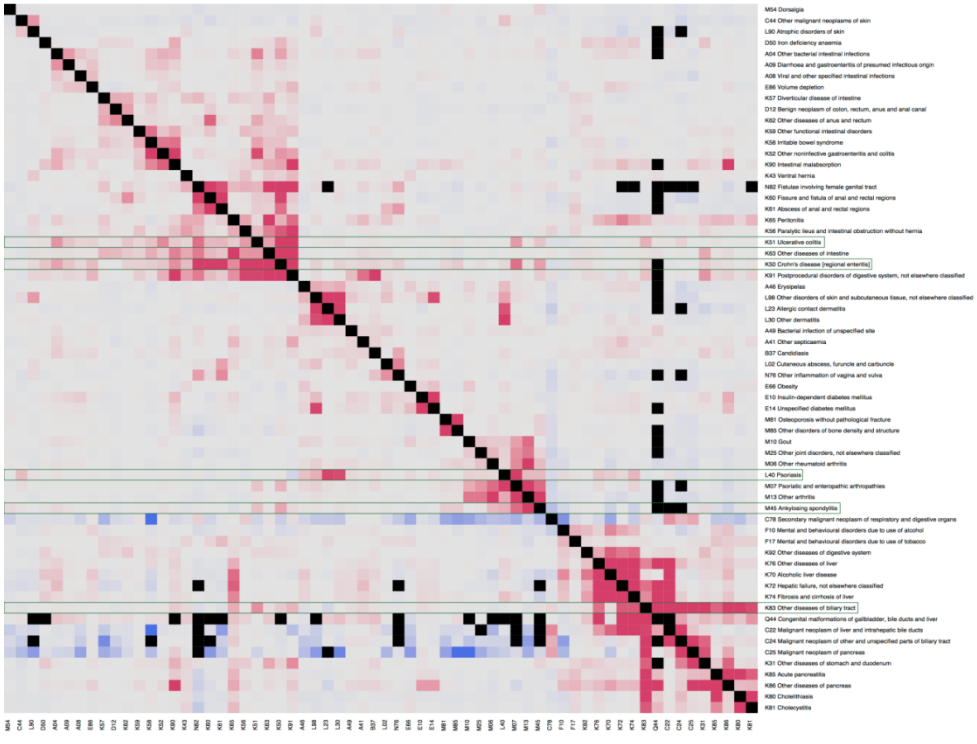
Temporal disease associations



A and B



A before B



Patient counts

Diagnoses		Patient count of order			Days from A and B		
A	B	A first	B first	Same	average	median	variance
Ankylosing spondylitis	Psoriasis	51	64	9	-538.47	-129.5	2250.31
	Sclerosing cholangitis	11	7	2	146.55	136.0	2006.63
	Ulcerative colitis	117	171	23	-658.93	-357.0	2395.10
	Crohn's disease	124	147	31	-305.49	0.0	2348.52
Psoriasis	Sclerosing cholangitis	31	4	0	2413.29	2475.0	2006.80
	Ulcerative colitis	134	120	7	133.35	41.0	2733.01
	Crohn's disease	77	122	4	-577.80	-769.0	2693.91
Sclerosing cholangitis	Ulcerative colitis	90	305	68	-1060.55	-272.0	2032.90
	Crohn's disease	58	124	21	-739.25	-456.0	2090.89
Ulcerative colitis	Crohn's disease	3523	1999	405	413.33	64.0	1854.68

Order of pairs of diagnoses for 7,334 patients with two or more diagnoses. The average and median number of days between the A and B diagnosis is also shown (negative indicates that B is before A.

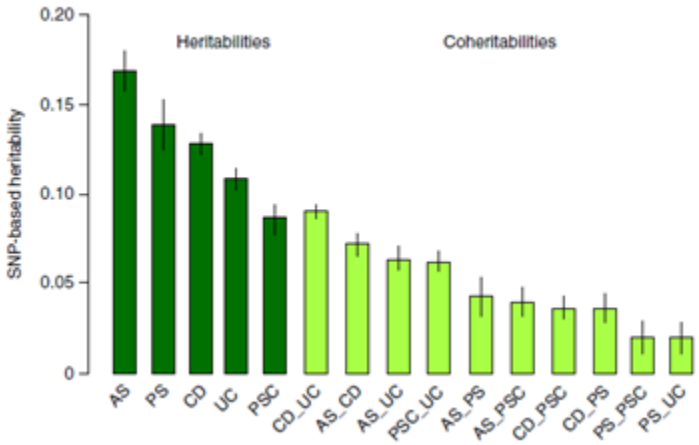
Shared or distinct genetic etiology?

Pleiotropy (sharing of risk alleles by disease A and disease B) or **Heterogeneity** (a subgroup of disease A cases has a higher load of risk alleles for disease B)?

Genetic and healthdata overlap analyses shows

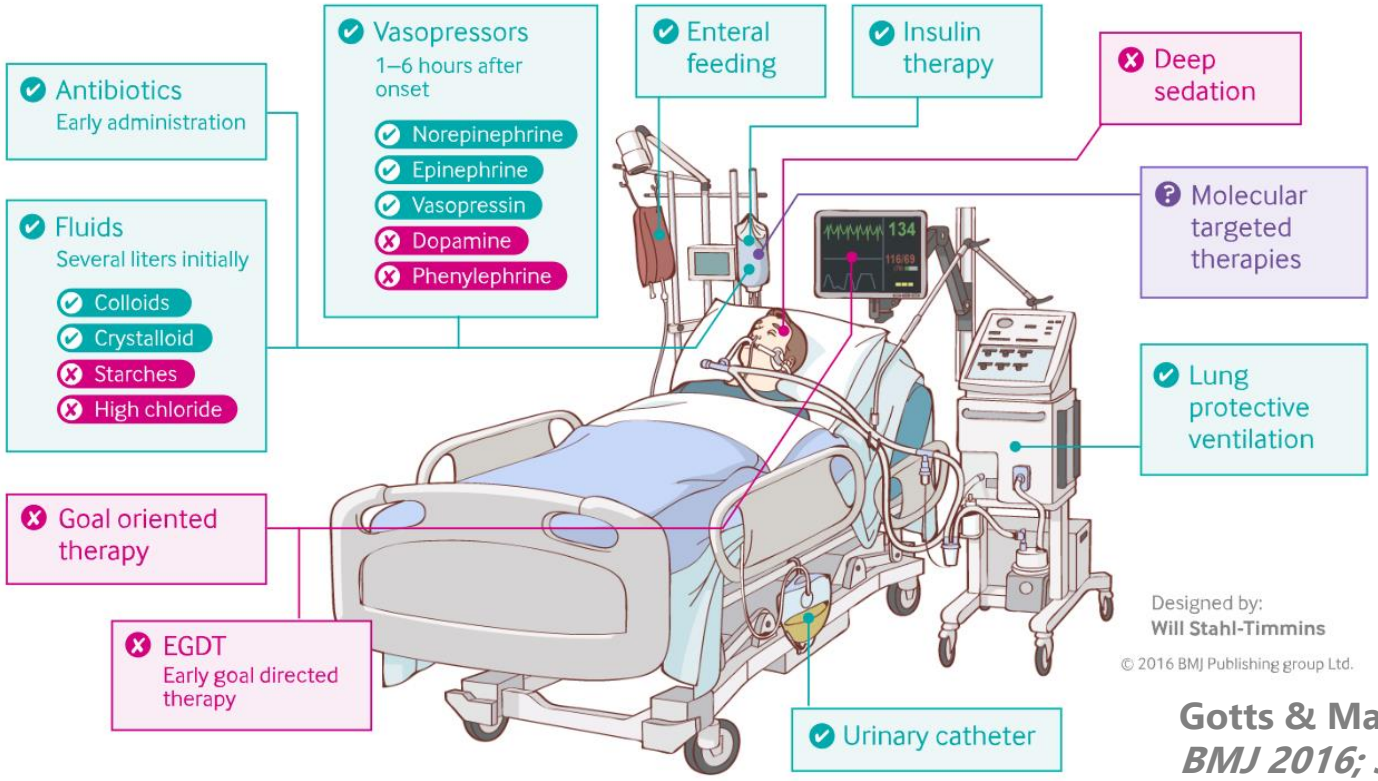
- **that shared pathophysiological pathways are the basis for clinical co-occurrence**
- that patients with concomitant syndromes are genetically distinct from patients without concomitant syndromes.

Immunochip-wide pleiotropy estimates

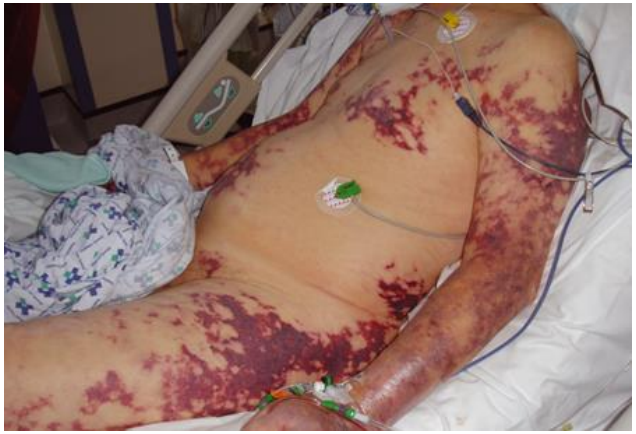


Ellinghaus et al. Nature Genetics, 48:510-518, 2016

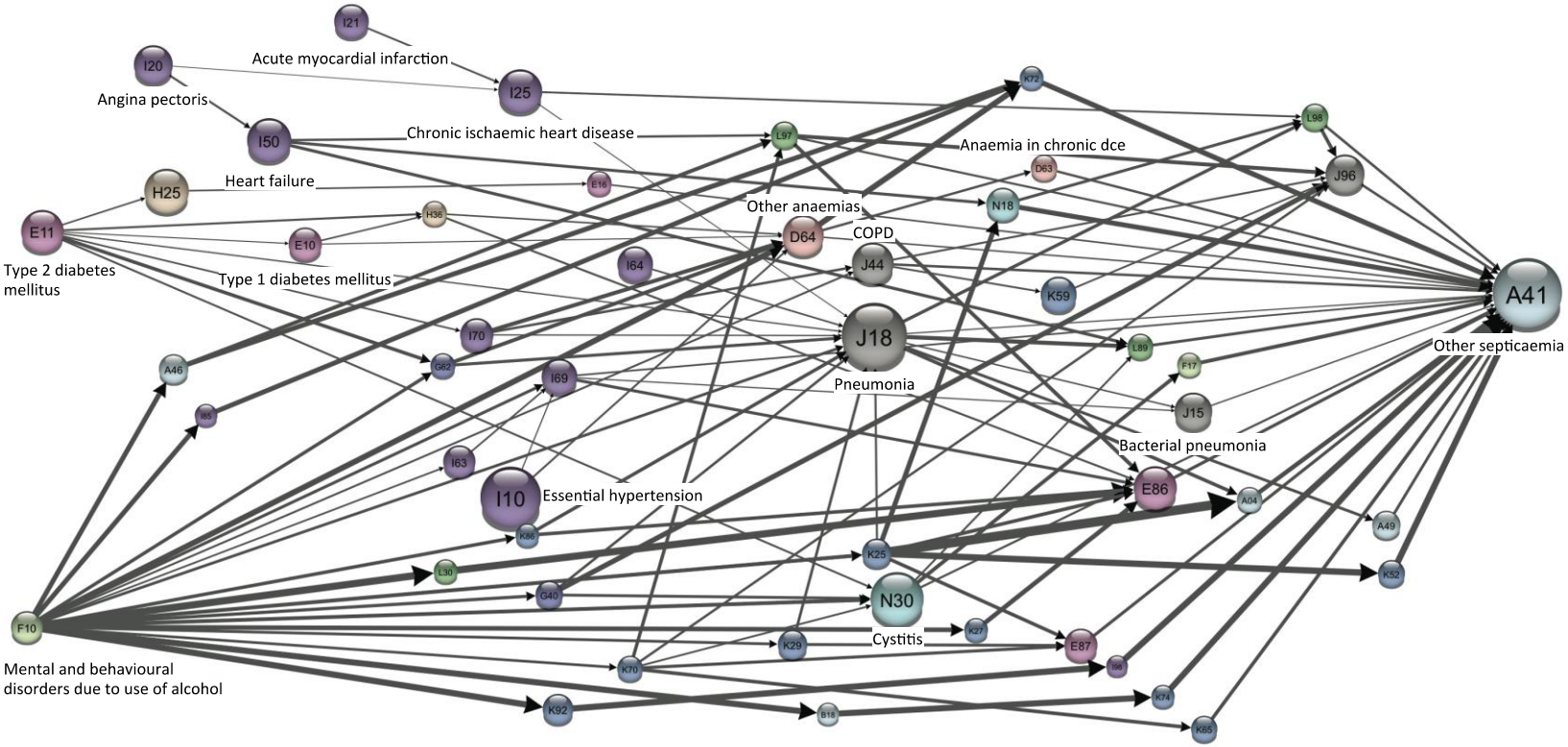
Sepsis



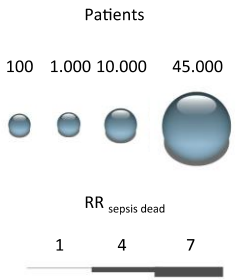
Gotts & Matthay
BMJ 2016; 353:i1585



Sepsis survival across pre-history
(120,000 patients, 56 significant trajectories)

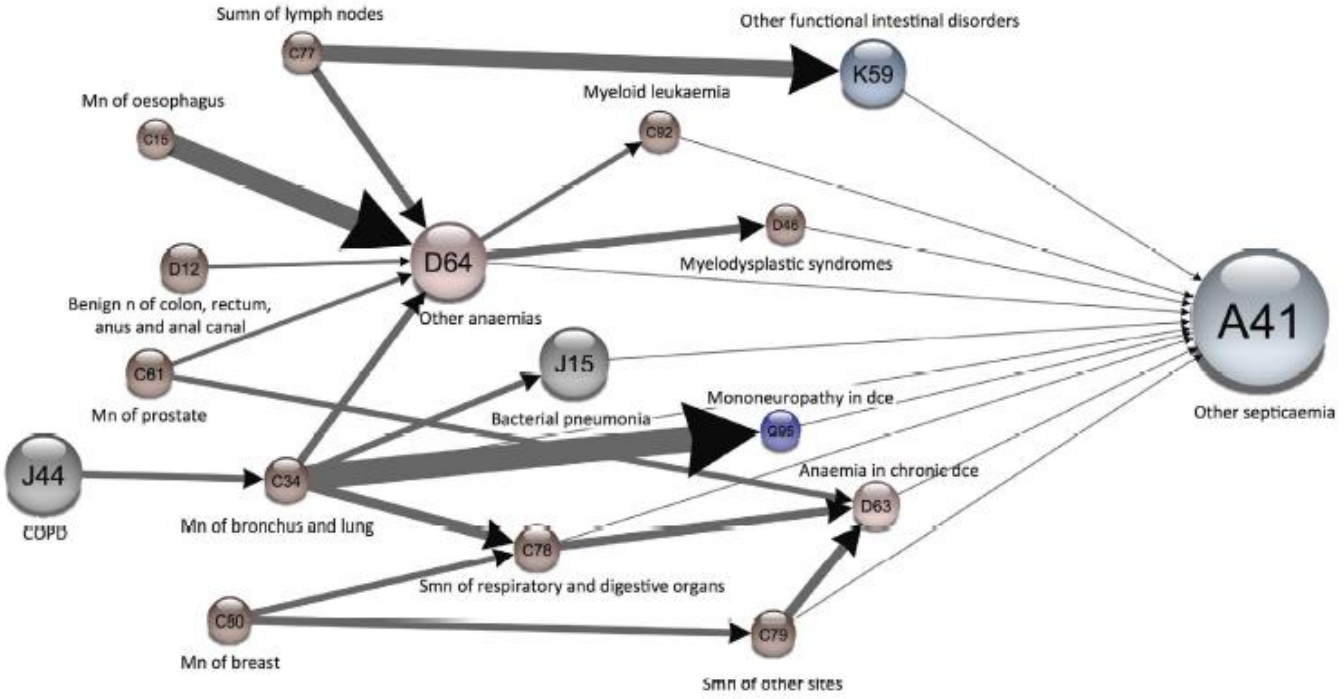


A04	Other bacterial intestinal infections
A46	Erysipelas
A49	Bacterial infection of unspecified site
B18	Chronic viral hepatitis
E16	Other disorders of pancreatic internal secretion
E87	Other disorders of fluid, electrolyte and acid-base balance
G40	Epilepsy
G62	Other polyneuropathies
H25	Senile cataract
H36	Retinal disorders in diseases classified elsewhere
I63	Cerebral infarction
I69	Sequelae of cerebrovascular disease
I70	Atherosclerosis
I85	Oesophageal varices
I98	Other disorders of circulatory system in diseases classified elsewhere
J96	Respiratory failure, not elsewhere classified
K25	Gastric ulcer
K27	Peptic ulcer, site unspecified
K29	Gastritis and duodenitis
K52	Other noninfective gastroenteritis and colitis
K59	Other functional intestinal disorders
K65	Peritonitis
K70	Alcoholic liver disease
K72	Hepatic failure, not elsewhere classified
K74	Fibrosis and cirrhosis of liver
K86	Other diseases of pancreas
K92	Other diseases of digestive system
L30	Other dermatitis
L89	Decubitus ulcer
L97	Ulcer of lower limb, not elsewhere classified
N18	Chronic renal failure



Beck et al, 2016

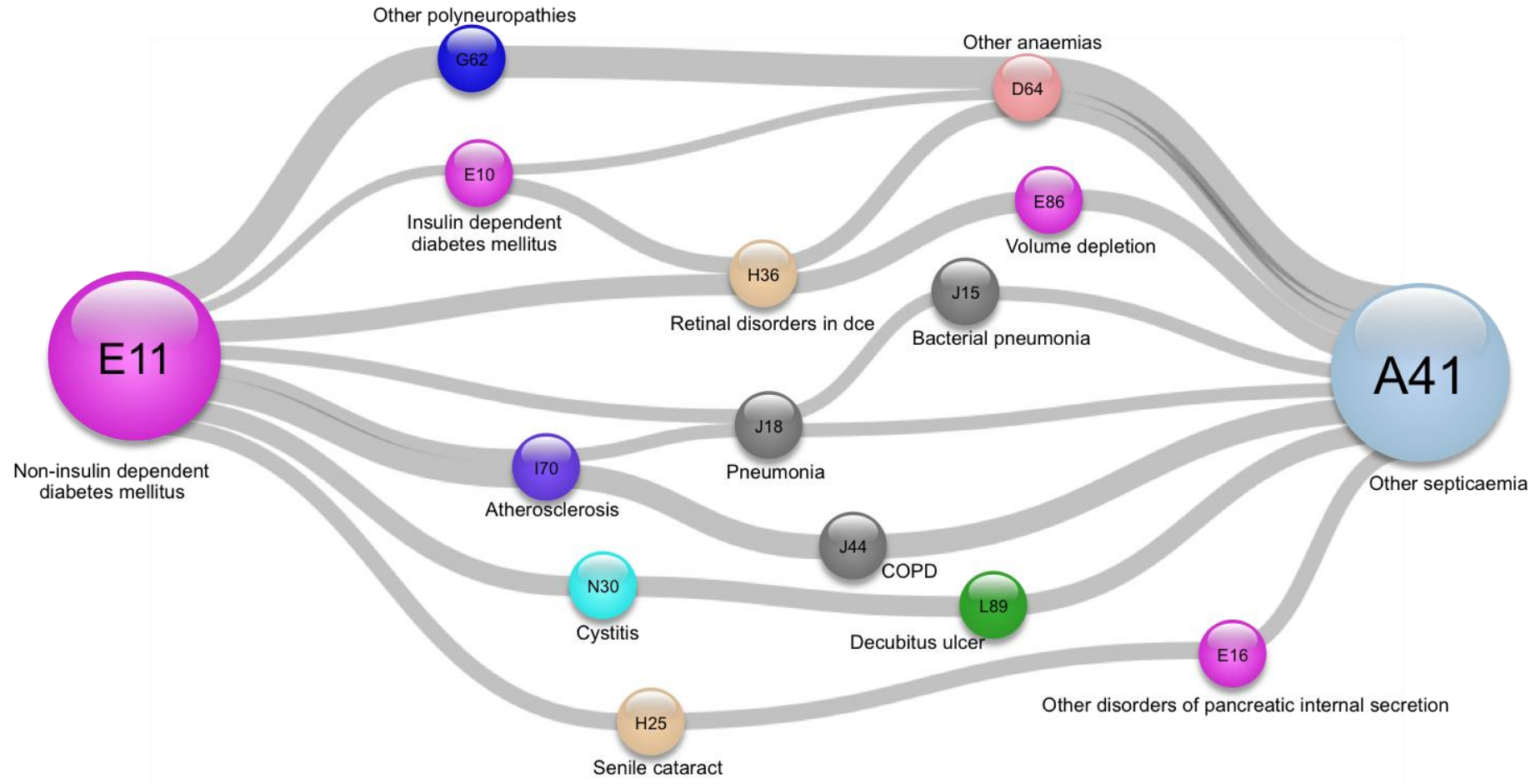
Cancer-subnetwork towards sepsis



Cancer-sepsis network from 16 trajectories that significantly altered RRsepsis dead. These length three sepsis-trajectories contain a minimum of one diagnosis from "Cancers (C00-C96)" in ICD-10 chapter 2: Neoplasms. Node size reflects the number of sepsis patients. Width of arrows indicates RRsepsis dead for a particular step in a trajectory.

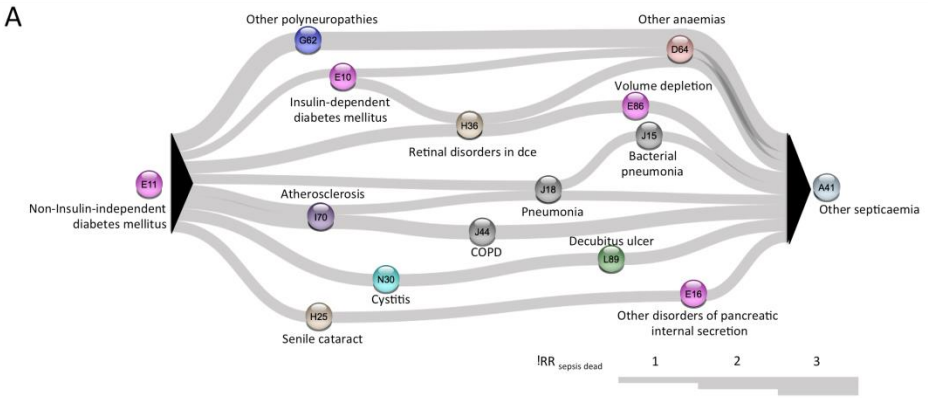
Beck et al, 2016

From diabetes to sepsis

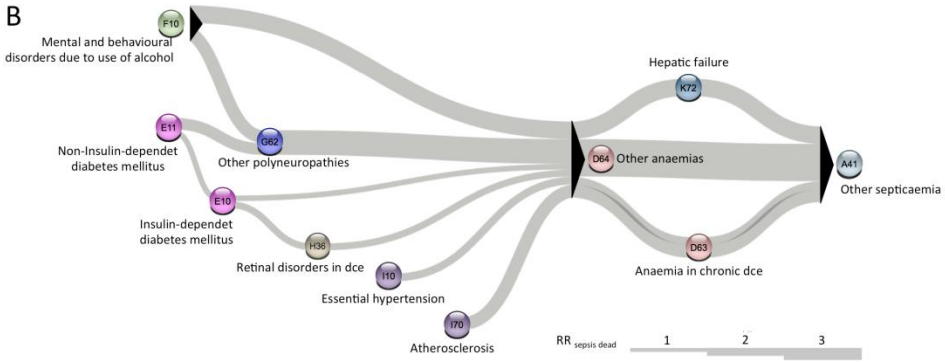


Beck et al, 2016

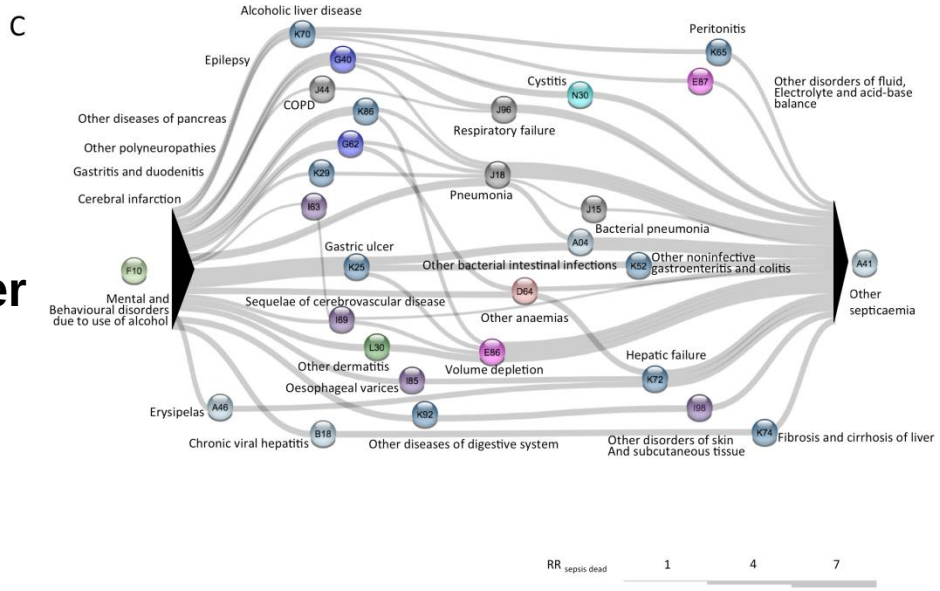
Diabetes

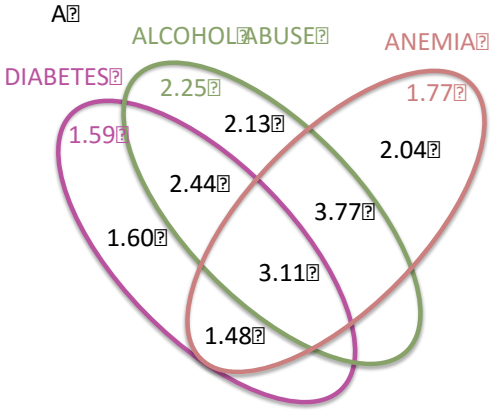


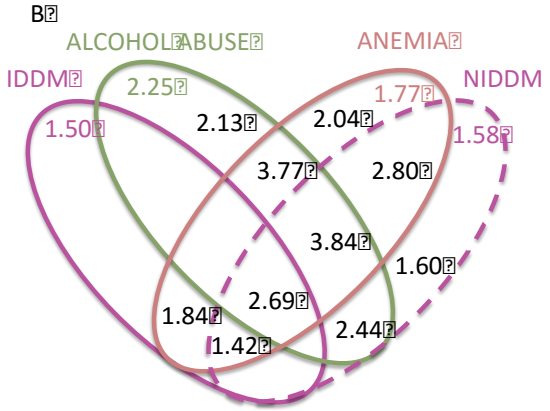
Anaemia



Mental disorder







The NEW ENGLAND JOURNAL of MEDICINE

CLINICAL IMPLICATIONS OF BASIC RESEARCH

A Wake-up Call for Type 2 Diabetes?

Shanta J. Persaud, Ph.D., and Peter M. Jones, Ph.D.

N Engl J Med 2016; 375:1090-1092 | September 15, 2016 | DOI: 10.1056/NEJMcibr1607950

This study examines an established association between a variant in a melatonin-receptor gene and type 2 diabetes, yielding insights into how the variant confers susceptibility to the disease.

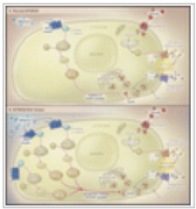
Disclosure forms provided by the authors are available at NEJM.org.

SOURCE INFORMATION

From the Diabetes Research Group, Division of Diabetes and Nutritional Sciences, King's College London, London.

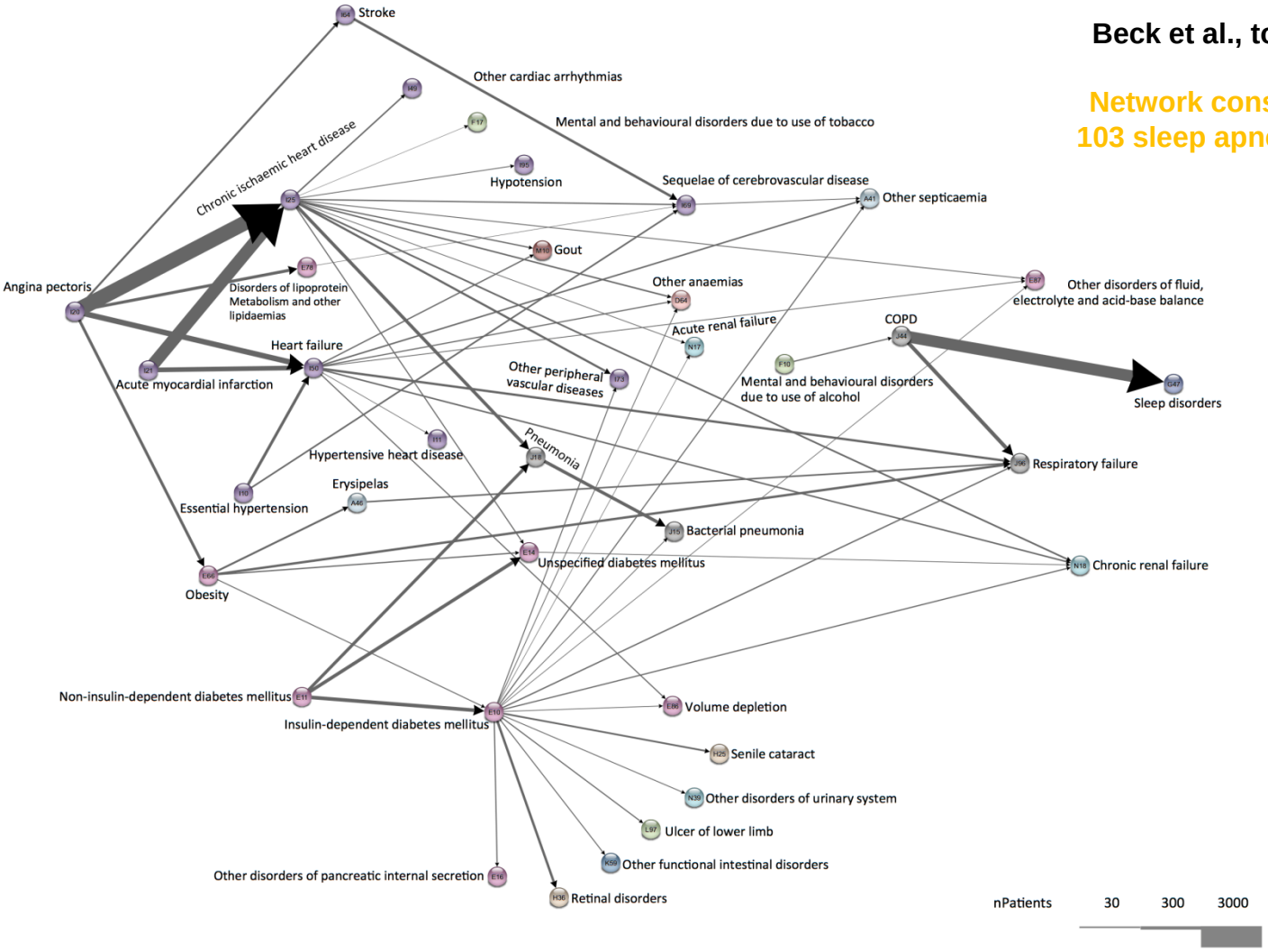
MEDIA IN THIS ARTICLE

FIGURE 1



Effect of Variant *MTNR1B* on Melatonin Signaling in Islet Beta Cells.

Sleep Apnea and Diabetes

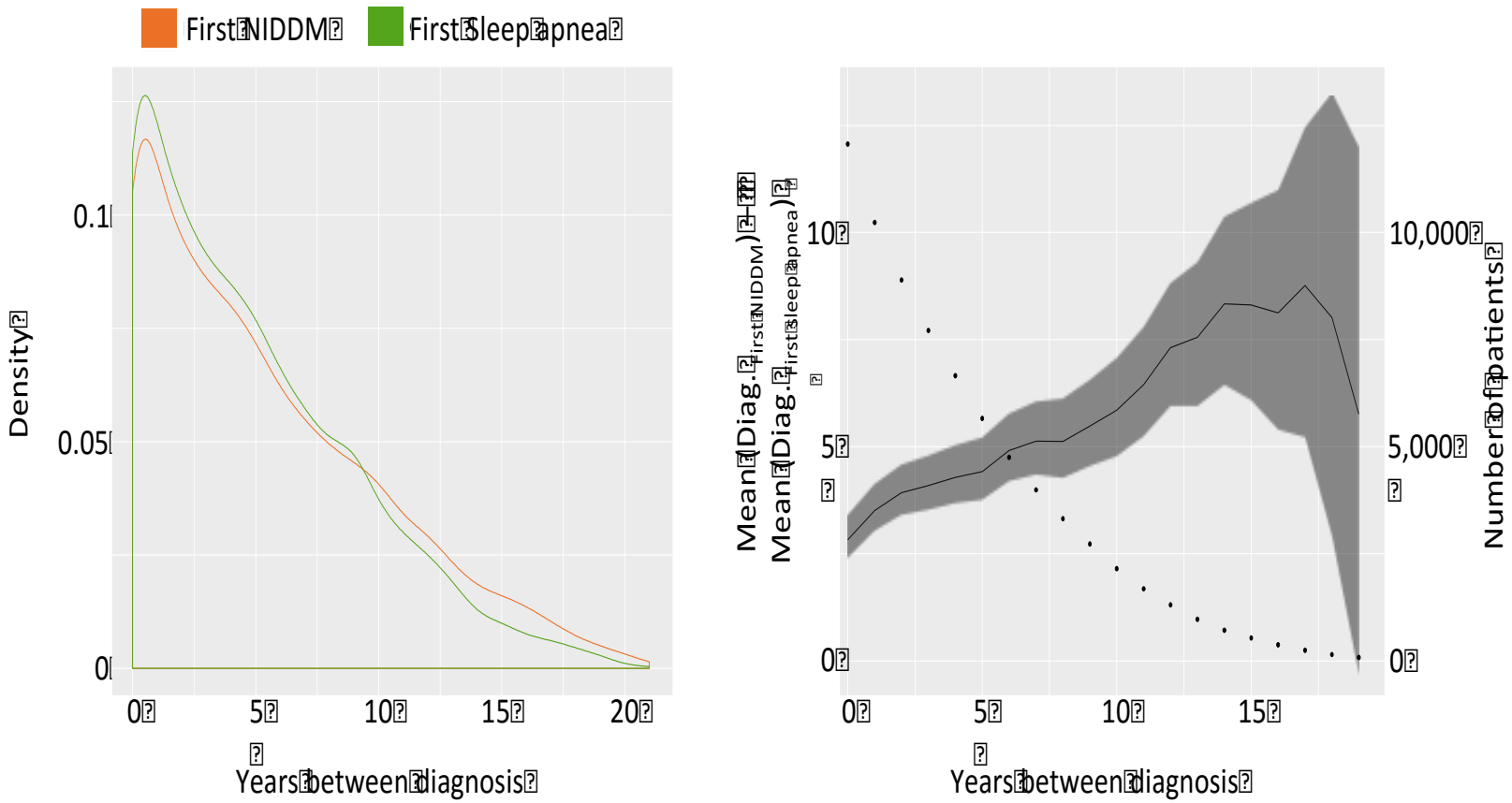


Beck et al., to appear 2017

Network constructed from 103 sleep apnea trajectories

No significant diagnosis direction, but patients diagnosed with diabetes before sleep apnea have significantly more comorbidities than patients diagnosed with sleep apnea before diabetes.

Excess number of comorbidities if diabetes is first



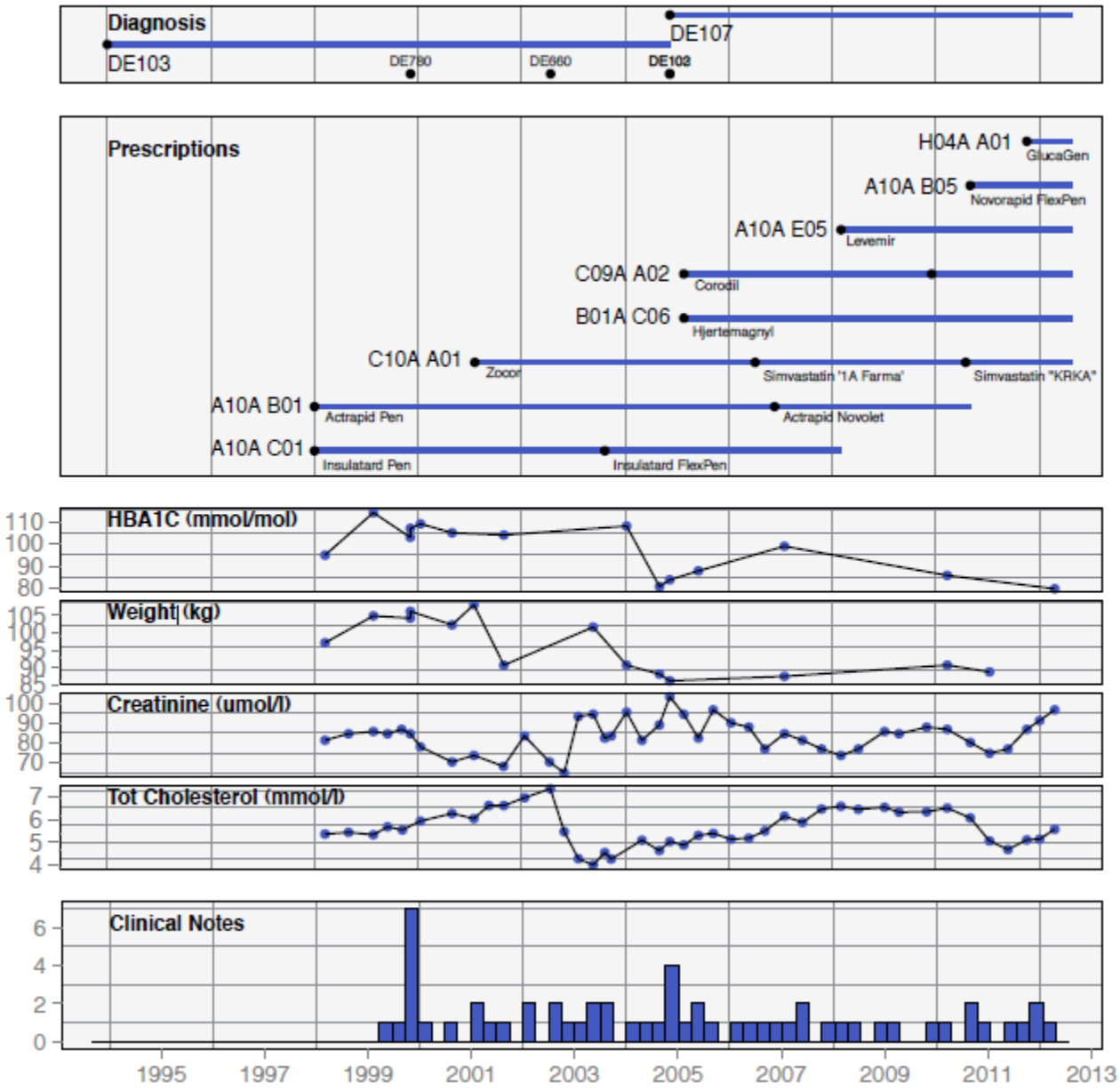
(A) Distribution of years between NIDDM and sleep apnea for patients diagnosed with NIDDM first (orange) and for patients diagnosed with sleep apnea first (green). (B) The plot shows the excess number of comorbidities for patients diagnosed with NIDDM first compared to those diagnosed with sleep apnea first (black line) with the 95% confidence interval (grey area). The dots indicate the number of patients having the minimum years between the two diagnoses.

Diagnoses

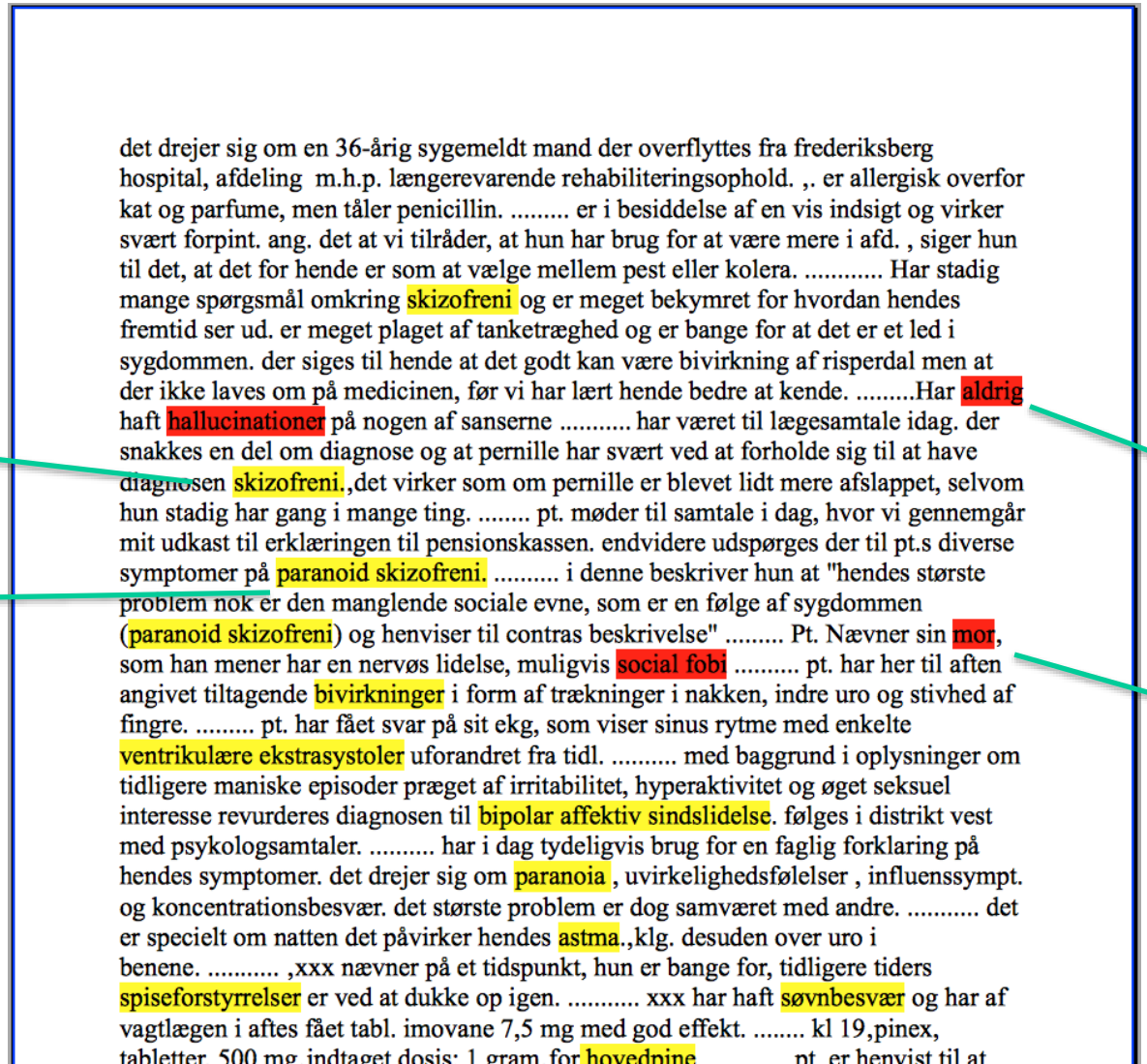
Drugs

Lab tests

Text



Deep-phenotyping by text mining of ICD10 terms in patient records



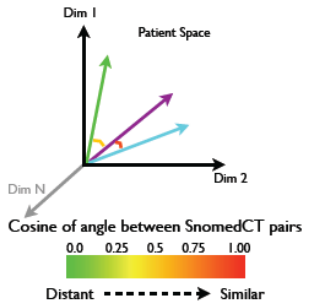
det drejer sig om en 36-årig sygemeldt mand der overflyttes fra frederiksberg hospital, afdeling m.h.p. længerevarende rehabiliteringsophold. ,. er allergisk overfor kat og parfume, men tåler penicillin. ..... er i besiddelse af en vis indsigt og virker svært forpint. ang. det at vi tilråder, at hun har brug for at være mere i afd. , siger hun til det, at det for hende er som at vælge mellem pest eller kolera. ..... Har stadig mange spørgsmål omkring skizofreni og er meget bekymret for hvordan hendes fremtid ser ud. er meget plaget af tanketrægghed og er bange for at det er et led i sygdommen. der siges til hende at det godt kan være bivirkning af risperdal men at der ikke laves om på medicinen, før vi har lært hende bedre at kende. .....Har aldrig haft hallucinationer på nogen af sanserne ..... har været til lægesamtale idag. der snakkes en del om diagnose og at pernille har svært ved at forholde sig til at have diagnosen skizofreni.,det virker som om pernille er blevet lidt mere afslappet, selvom hun stadig har gang i mange ting. ..... pt. møder til samtale i dag, hvor vi gennemgår mit udkast til erklæringen til pensionskassen. endvidere udspørges der til pt.s diverse symptomer på paranoid skizofreni. ..... i denne beskriver hun at "hendes største problem nok er den manglende sociale evne, som er en følge af sygdommen (paranoid skizofreni) og henviser til contras beskrivelse" ..... Pt. Nævner sin mor, som han mener har en nervøs lidelse, muligvis social fobi ..... pt. har her til aften angivet tiltagende bivirkninger i form af trækninger i nakken, indre uro og stivhed af fingre. ..... pt. har fået svar på sit ekg, som viser sinus rytme med enkelte ventrikulære ekstrasystoler uforandret fra tidl. ..... med baggrund i oplysninger om tidligere maniske episoder præget af irritabilitet, hyperaktivitet og øget seksuel interesse revurderes diagnosen til bipolar affektiv sindslidelse. følges i distrikt vest med psykologsamtaler. ..... har i dag tydeligvis brug for en faglig forklaring på hendes symptomer. det drejer sig om paranoia , uvirkelighedsfølelser , influenssympt. og koncentrationsbesvær. det største problem er dog samværet med andre. ..... det er specielt om natten det påvirker hendes astma.,klg. desuden over uro i benene. ..... ,xxx nævner på et tidspunkt, hun er bange for, tidligere tiders spiseforstyrrelser er ved at dukke op igen. ..... xxx har haft søvnbesvær og har af vagtlægen i aftes fået tabl. imovane 7,5 mg med god effekt. ..... kl 19,pinex, tabletter, 500 mg, indtaget dosis: 1 gram, for hovedpine. pt. er henvist til at

F20

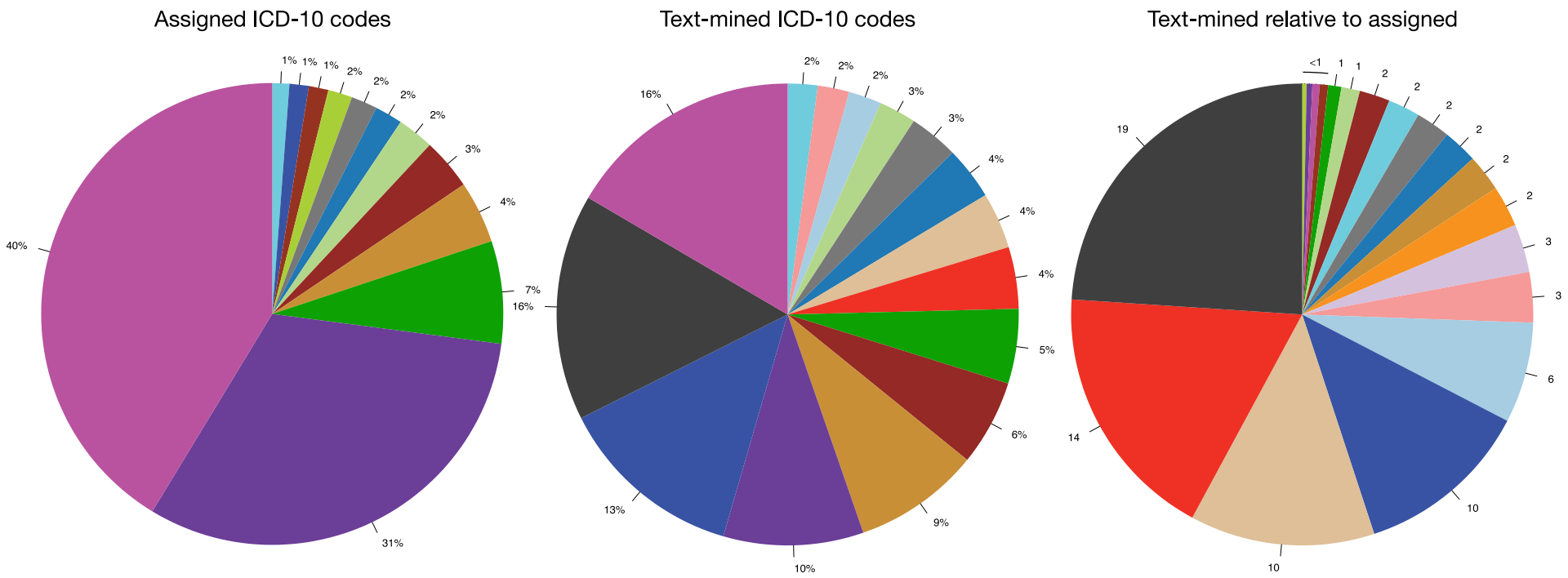
F200

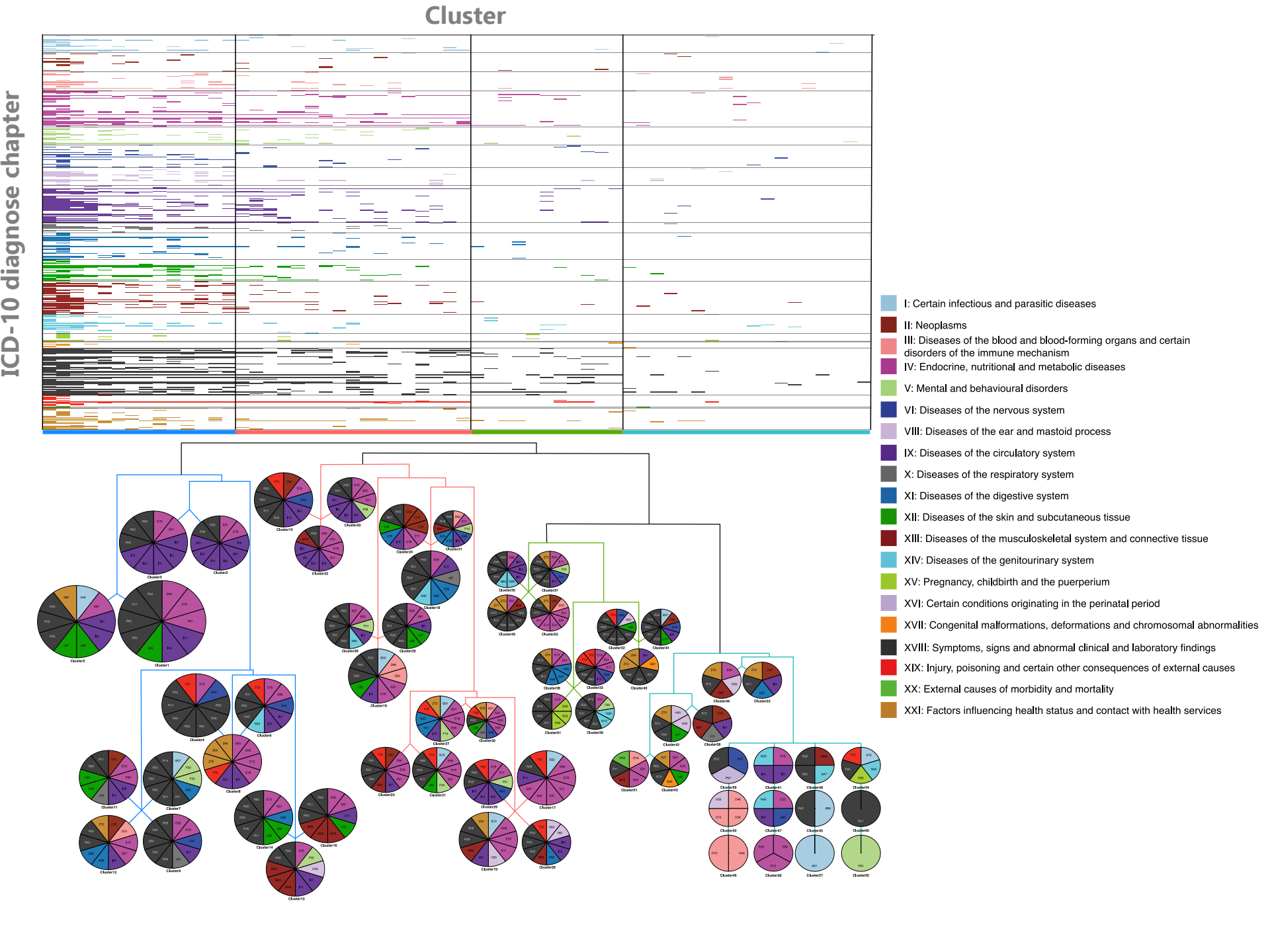
Negation

Family

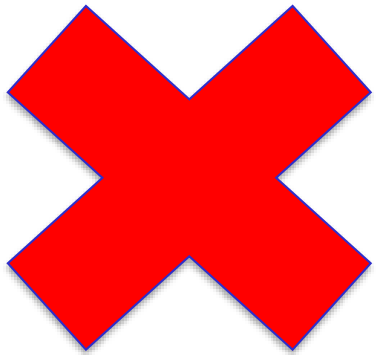


Distribution of assigned and text-mined codes, Steno Diabetes Center




Cluster
ICD-10 diagnose chapter
I: Certain infectious and parasitic diseases
II: Neoplasms
III: Diseases of the blood and blood-forming organs and certain disorders of the immune mechanism
IV: Endocrine, nutritional and metabolic diseases
V: Mental and behavioural disorders
VI: Diseases of the nervous system
VIII: Diseases of the ear and mastoid process
IX: Diseases of the circulatory system
X: Diseases of the respiratory system
XI: Diseases of the digestive system
XII: Diseases of the skin and subcutaneous tissue
XIII: Diseases of the musculoskeletal system and connective tissue
XIV: Diseases of the genitourinary system
XV: Pregnancy, childbirth and the puerperium
XVI: Certain conditions originating in the perinatal period
XVII: Congenital malformations, deformations and chromosomal abnormalities
XVIII: Symptoms, signs and abnormal clinical and laboratory findings
XIX: Injury, poisoning and certain other consequences of external causes
XX: External causes of morbidity and mortality
XXI: Factors influencing health status and contact with health services

Which disease-disease and symptom correlations are treatment related?



Text mining of drug names, ADE/ADRs, diagnoses, …

Removed ADR - no corresponding structured data

Behandlet med Zyprexa 5 mg fra 3. til 24.6.99 og 10 mg fra 24. til 29.6.99 med nogen effekt på tankeforstyrrelser, men seponeret pga appetitøgning. Herefter Risperdal 2 mg stigende til 4 mg i perioden 29.6. til 12.7.99, men seponeret på grund af uro i kroppen og "ostek-lokkefornemmelse". Herefter Orap 2 mg fra 2.8. stigende til 3 mg fra 30.8.99 med god effekt på tankeekko og tankemylder. Behandlet med Zoloft 50 mg fra maj 98 til maj 99 med noget virkning på depressive symptomer, men seponeret på grund af natlig svedtendens. Siden 14.7.99 Efexor 75 mg med nogen effekt på antallet og sværhedsgraden af kortvarige depressive episoder

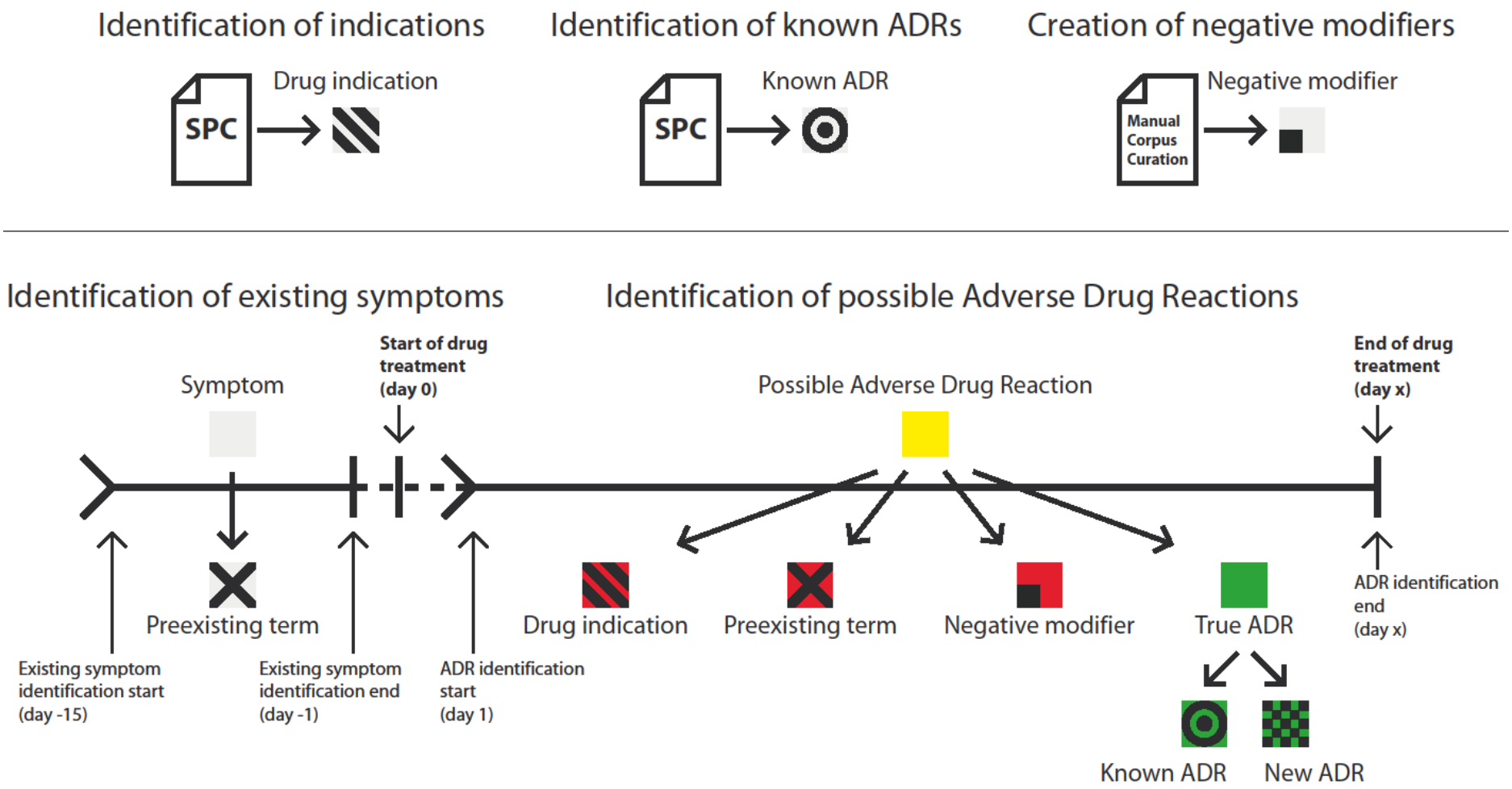
se venligst under allergier. Desuden forsøgt beh med Zyprexa, sep grundet vægtøgning, træthed og manglende effekt. Risperdal ord med nogen effekt tillagt dogmatil(1999). Efterfølgende aurorix beh seponeret i 1999. Startede istedet remeron. Aktuel medicindois, jvf ud-skrivningsnotat fra U12 samt EPJmedicinliste.

tbl. leponex 100+0+0+200 mg,tbl. rivotril 0,5 +0+1+ 0 mg,tbl. arintapin 0+0+0+30 mg,tbl. klomipramin 0+0+0+25 mg,tbl. imoclone 7,5 mg nocte,tbl. rivotril 2 mg p.n. max x 1 dgl. tbl. marevan a 2,5 mg efter skema,tbl. magnesia 1 g p.n. laxoberaldr. 7,5 mg /ml 15 dr. p.n. mix. link 150 mg /ml 15 ml p.n. max x 3 dgl. Figensaft 20+0+20+0 ml,Pt. er aktuelt,CAVE,tricykliske antidepressiva. Dette kan dog ikke bekræftes og pt. har tidl. fået imipramin, som han har tålt godt, hvorfor der er ansøgt om ophævelser af denne cave på højere niveau. Har tidl. fået zyprexa som blev sep. grundet vægtøgning, træthed og manglende effekt

Removed ADR - negation and subject identification

...Jeg mener fortsat, at han har brug for medicin, da han i går fx var meget vred og følte sig utryg og angst og har haft svært ved at sove. Dette synes pt. at accepterer. Jeg tilbyder herefter Zyprexa i stedet for Risperdal, pt. afviser dette, da han ved medpatient har fået denne medicin og har fået øget appetit, dette vil han ikke. Har ikke tidligere fået antipsykotisk medicin. Accepterer herefter Cisordinol, startende på en lille dosis. Accepterer også angstdæmpende medicin i dagtiden. Angiver, at når han bliver vred kan han godt styre det. Synes det hjalp i går noget at få Nozinan. Virker fortsat garderet. Siger intet uopfordret. Sparsomt sygdomsindsigt. Er ikke vredladen. Er i dag et heller latent aggressiv...

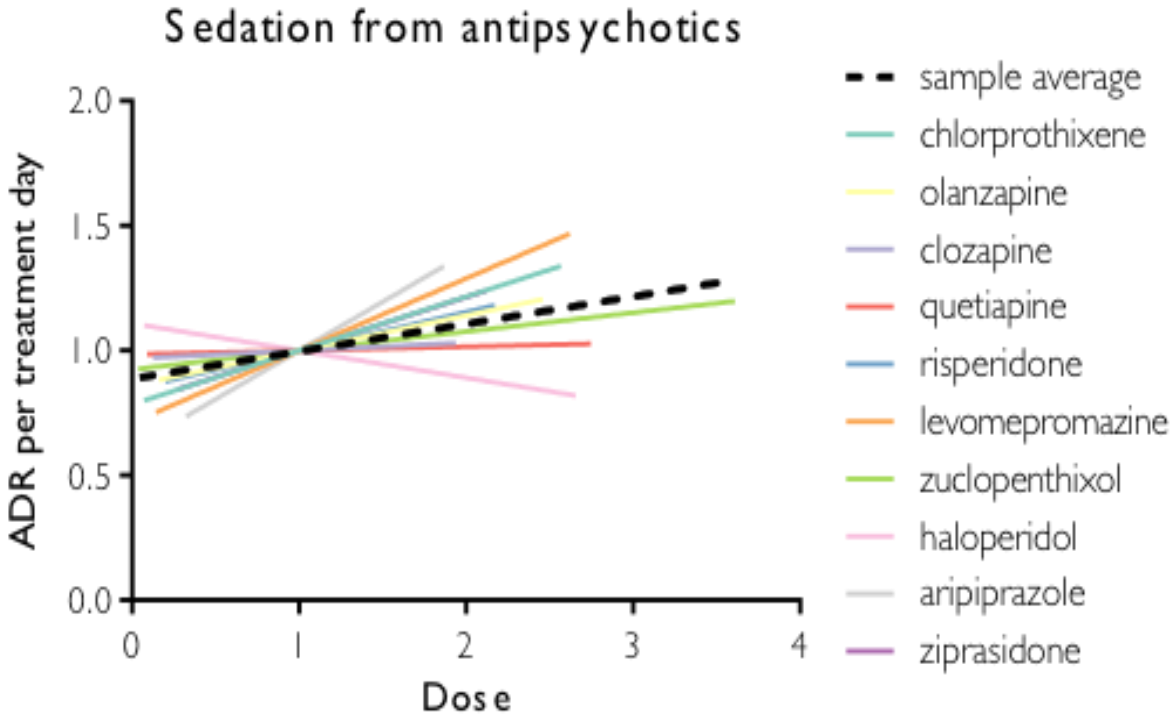
Text mining Adverse Drug Reactions
(using 7,500 drug names and 21,000 ADRs)



Eriksson et al. Drug Safety 2014

ADR-dose dependencies

Dosages from structured medication data



ADRs and doses are normalized on multiples of the minimum dose prescribed of each drug.

Plot for 21 days steady dosage data is visualized, sample average slope 0.1105 (95% CI, 0.03085-0.1901), non-zero slope p-value was 0,0074, all individual drug slopes are positive except for haloperidol.

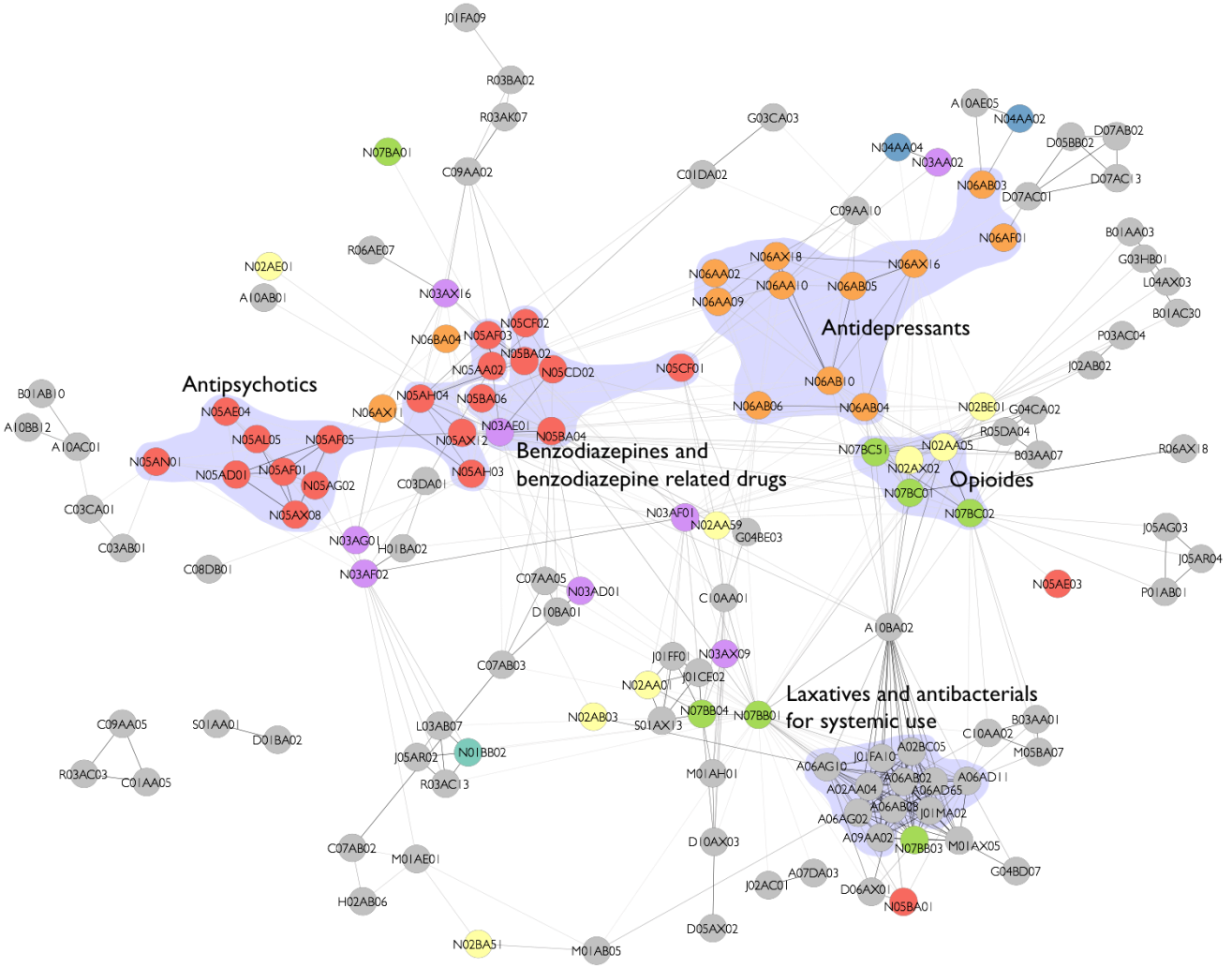
Eriksson et al. Drug Safety 2014

Drug-ADR similarities

(ADRs text mined by temporal analysis of EHRs)

- N01 Anesthetics
- N02 Analgesics
- N03 Antiepileptics
- N04 Anti-parkinson drugs
- N05 Psycholeptics
- N06 Psychoanaleptics
- N07 Other nervous system drugs
- All other ATC codes

13 years of drug use at Mental centre Sct. Hans ATC code coloring. Edges show Drug ADR profile similarity, darker edge indicates stronger similarity. Network contains 500 strongest edges (Jaccard index). "N03 Antiepileptics" are scattered, not showing any clustering. This is expected as antiepileptics is a very diverse drug classein terms of ADRs. Laxatives and antibacterials for systemic use group since both cause diarrhea and stomach ache and other gastrointestinal problems



Possible ADRs?

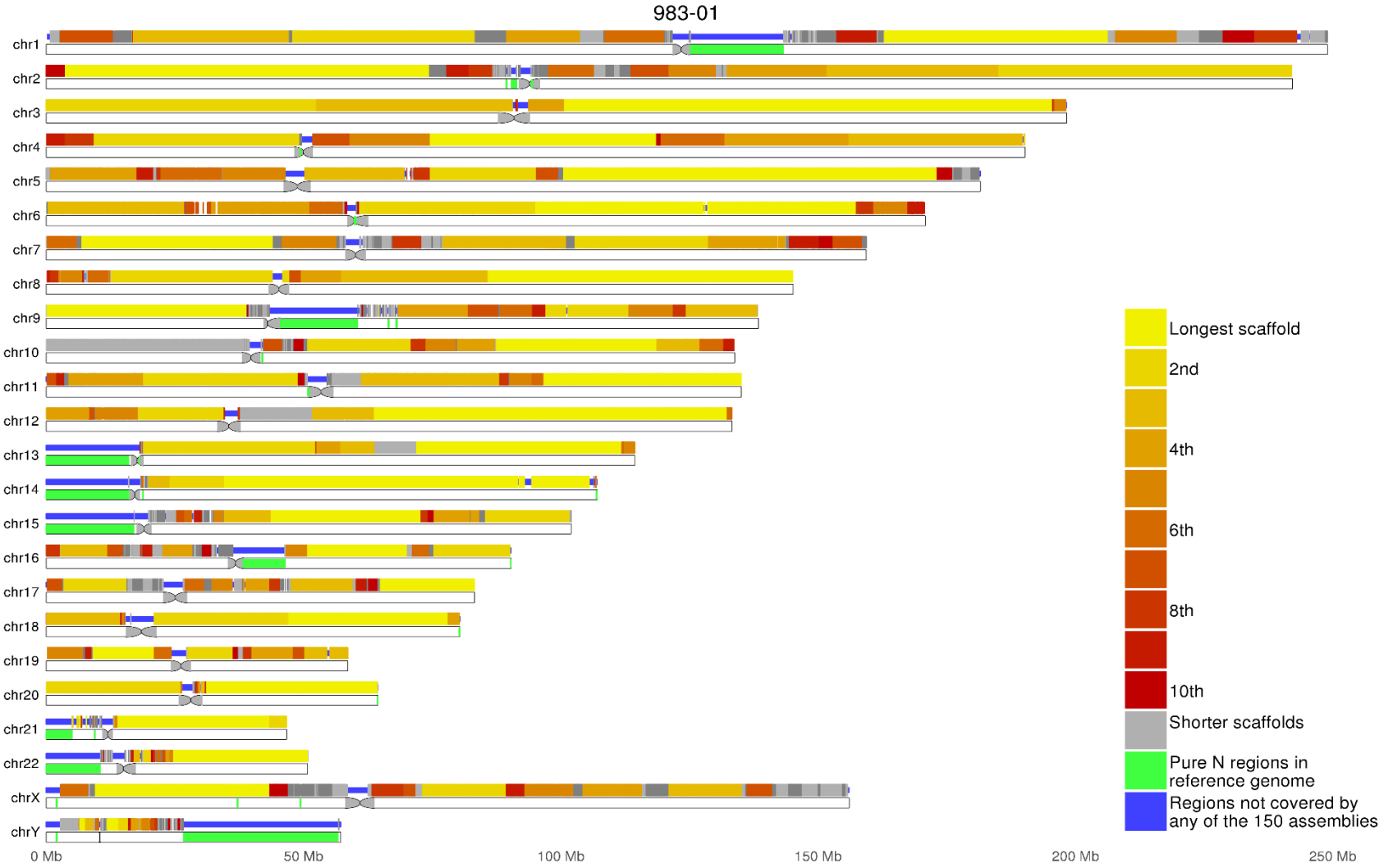
Drug substance	ADE	p-value
Dipyridamole	Visual impairment	4.375e-04
Simvastatin	Personality changes	8.408e-08
Citalopram	Psychosis	8.807e-04
Bendroflumethiazide	Apoplexy	8.46e-03
Chlordiazepoxide	Nystagmus	4.03e-08

X1	Y1	p1
X2	Y2	p2
....		

p-values are multiple testing corrected

Trio-based Danish reference genome finished

Individual high quality genome shown on the official reference for the humane genome



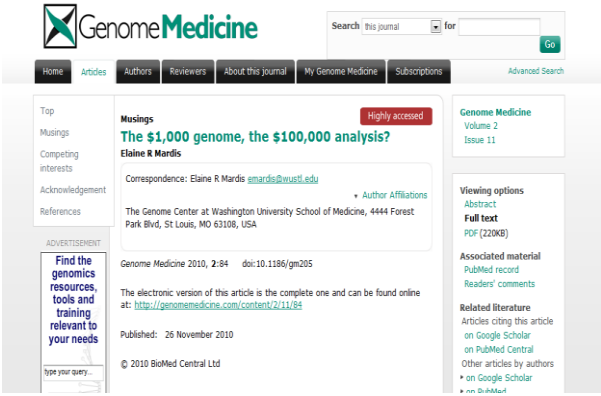
Personlig Medicin og Individualiseret Behandling

Oplæg til en samlet dansk indsats

22. juni 2015

Region Sjælland – vi er til for dig
Syddansk Universitet
Aalborg Universitet – Ad nye veje
Danske Patienter
Aarhus Universitet
Danske Regioner
Region Hovedstaden
Region Nordjylland
Region Syddanmark
midt regionmidtjylland

Δ population health data



- **Health data driven:**
  - **Redefine phenotypes as trajectories**
  - **Enable prediction using predictable trajectories?**
  - **Handle noise better**
  - **Handle life long data capture**
  - **”Live data” versus data dumps versus registers**

- **Include what is not in the patient records in new ways:**
  - **Diet,**
  - **Income, …**
  - **Education, grades in exams, …**

Acknowledgements

EPR and registry analysis

Anders Bøck Jensen, CPR/KU
Isa Kirk, CPR/KU
Christian Simon, CPR/KU
Karina Basanik, CPR/KU
Mette Beck, CPR/KU
Peter Bjødstrup Jensen, CPR/KU
Robert Eriksson, CPR/KU
Pope Mosely, U. New Mexico
Tudor Oprea, U. New Mexico
Lars Juhl Jensen, CPR/KU
Francisco Simões Roque, CBS/DTU
Eva Roitmann, CBS/DTU
Karen Søeby, Hvidovre Hospital
Søren Bredkjær, Region Sealand

Mental Centre Sct. Hans

Thomas Werge
Henriette Schmock
Thomas Hansen, Sct. Hans Hospital

Steno Diabetes Center, Hagedorn, Novo Nordisk A/S

Peter Rossing
Henrik Ullits Andersen
Regine Bergholdt
Thomas Almdal, now Gentofte Hospital
Flemming Pociot, now Herlev Hospital
Torben Hansen, now U. Copenhagen
Oluf Borbye Pedersen, now U. Copenhagen

Nature Reviews Genetics | AOP, published online 2 May 2012; doi:10.1038/nrg3208

TRANSLATIONAL GENETICS

Mining electronic health records: towards better research applications and clinical care

Peter B. Jensen[1], *Lars J. Jensen*[1] *and Søren Brunak*[1,2]

Abstract | Clinical data describing the phenotypes and treatment of patients represents an underused data source that has much greater research potential than is currently realized. Mining of electronic health records (EHRs) has the potential for establishing new patient-stratification principles and for revealing unknown disease correlations. Integrating EHR data with genetic data will also give a finer understanding of genotype–phenotype relationships. However, a broad range of ethical, legal and technical reasons currently hinder the systematic deposition of these data in EHRs and their mining. Here, we consider the potential for furthering medical research and clinical care using EHR data and the challenges that must be overcome before this is a reality.

Clinical decision support
(CDS). Software systems providing support for decision making to physicians through the application of health knowledge and logical rules to patient data.

Biobanks
Central repositories of biological material that are mainly used for research. They facilitate the re-use of collected samples in different research projects.

Information technology has transformed the way health care is carried out and documented. Presently, the practice of health care generates, exchanges and stores huge amounts of patient-specific information. In addition to the traditional clinical narrative, databases in modern health centres automatically capture structured data relating to all aspects of care, including diagnosis, medication, laboratory test results and radiological imaging data.

This transformation holds great promise for the individual patient as richer information, coupled with clinical decision support (CDS) systems, becomes readily available at the bedside to support informed decision making and to improve patient safety[1,2]. ... especially interesting when traditional health-care-sector data is linked with biobanks and genetic data[4].

Despite the great potential, researchers who wish to analyse large amounts of patient data are still faced with technical challenges of integrating scattered, heterogeneous data, in addition to ethical and legal obstacles that limit access to the data[5,6]. It is hoped that large-scale adoption of health information technology (HIT) infrastructure in the form of electronic health records (EHRs) and agreed standards for interoperability and schemes for privacy and consent, will improve this situation (TABLE 1). With incentives for improved public health and the expected health budget savings[7,8], these matters