Payment Reform in the Era of Advanced Diagnostics, Artificial Intelligence and Machine Learning

James Sorace MD MS
jamessorace1@gmail.com

Overview

- Unique alignment of three trends/forces for change:
  1. Landmark 2015 Institute of Medicine report "Improving Diagnosis in Health Care".
  2. Ongoing scientific revolution that includes new diagnostic and therapeutic technologies as well as advances in AI/ML and health information technology.
  3. Tremendous interest in new advanced payment models (APMs) that improve health care quality while reducing cost.
- Current efforts at payment reform and HIT have met limited success.
- The concept of complex patients and its implications.
- New APMs based on data and knowledge sharing are needed.

Health Reform and Payment Models 1

- The need to innovate new payment models resulted in the Centers for Medicare and Medicaid Innovation (Affordable Care Act 2011).
  - Funded with $10 billion over 10 years. Over $5.6 billion obligated.
- Models that maintain or improve quality while reducing cost can be nationally implemented with the approval of the Actuary of CMS. **No other action is required**.
  - To date, according to the CMMI website 2 of 3 models advanced to the actuary have been certified.

CMMI Results

- It has proven difficult to achieve improvements.
  - GAO reported in May of 2018 that 4 of 37 Advanced Payment Models reduced cost and increased quality.
  - NEJM study Evaluation of Medicare's Bundled Payments Initiative for Medical Conditions: "Hospital participation in five common medical bundles under BPCI was not associated with significant changes in Medicare payments, clinical complexity, length of stay, emergency department use, hospital readmission, or mortality."
  - See here for the results of the Oncology Care Model.
- A recent review has highlighted the challenges faced by CMMI model developers, and recommendations in the following areas:
  - Iterative testing with market feedback.
  - Realistic time frames.
  - Model Integration.

Issues With Clinical Measures

- Must address the issues found in clinical measures.
  - Time Out - Charting a Path for Improving Performance Measurement**:**
    - 63% of physicians report that current quality measures do not capture the quality of the care they provide.
    - Physician practices estimated to spend $15.4 billion annually to report measures.
    - Review of 86 measures relevant for an ambulatory medicine practice on the 2017 QPP list: 37% were vailed, 35% were invalid, and 28% were of uncertain validity.
  - Relationship of primary care physicians' patient caseload with measurement of quality and cost performance:
    - "Relatively few primary care physician practices are large enough to reliably measure 10% relative differences in common measures of quality and cost performance among fee-for-service Medicare patients."
- These findings foreshadow a larger issue.

The Path Ahead

- Current approaches are based on top down centralized planning and measurement.
- New approaches needed that:
  - Question underlying assumptions about the complexity of disease processes and clinical measurement.
    - **No average patient!**
  - Focuses on the importance of timely diagnosis and correct patient categorization.
  - Emphasizes large scale decentralized data sharing and collaboration.
  - Rewards the appropriate use of AI/ML.
  - Improves clinical communication

How Complex Are Patients?

- To date efforts at health reform have fundamentally underestimated the complexity or variation of human disease.
- Two studies that quantitate this complexity with different approaches.
  - Medicare Disease Combination analysis.
  - Medicare Twin Study.
- Will demonstrate that there is no average or typical patient.
- Explains why current top down approaches are burdensome and can never make a significant impact.

Disease Combination Analysis

- Data included all 2008 Beneficiaries with continuous fee for service claims history.
  - 32,220,634 Beneficiaries
  - $283,088,306,347
- For more information on HCCs see: Risk adjustment of Medicare capitation payments using the CMS-HCC model. Pope GC, Kautter J, Ellis RP, Ash AS, Ayanian JZ, Lezzoni LI, Ingber MJ, Levy JM, Robst J. *Health Care Financing Review*. 2004 Summer; 25(4):119-41.
  - At the 184 CC level over 23 million DCs were detected. 99.6% of them were unique (contained only 1 beneficiary). To complex to interpret.
  - At the 70 HCC level 2,027,394 Disease Combinations (DCs) were detected.

Disease Combination Analysis

Four Groups Were Identified:

Group	% of Beneficiaries	% of Expenditures
1) No HCC	35	6
2) 100 most prevalent DCs	33	15
3) Remaining 2,072294 DCs	32	79
4) 1,658,233 Unique DCs	5.1	35

Example DCs by Prevalence (1-5 and 96-100)

DC Rank	Number of Beneficiaries (%)	HCC(s) describing the DC
1	1,667,891 (5.17647)	19_Diabetes without Complication
2	764,522 (2.37277)	10_Breast, Prostate, Colorectal and Other Cancer
3	723,760 (2.24626)	108_Chronic Obstructive Pulmonary Disease
4	610,943 (1.89612)	105_Peripheral Vascular Disease
5	531,536 (1.64968)	92_Specified Heart Arrhythmias
96	19,237 (0.05970)	27_Chronic Hepatitis
97	19,196 (0.05958)	54_Schizophrenia & 108_Chronic Obstructive Pulmonary Disease
98	18,806 (0.05837)	80_Congestive Heart Failure & 92_Specified Heart Arrhythmias & 131_Renal Failure
99	18,754 (0.05820)	101_Cerebral Palsy, Other Paralytic Syndromes
100	18,643 (0.05786)	38_Rheum Arthritis and Inflammatory Connective Tissue Disease & 55_Major Depressive, Bipolar, Paranoid Disorders

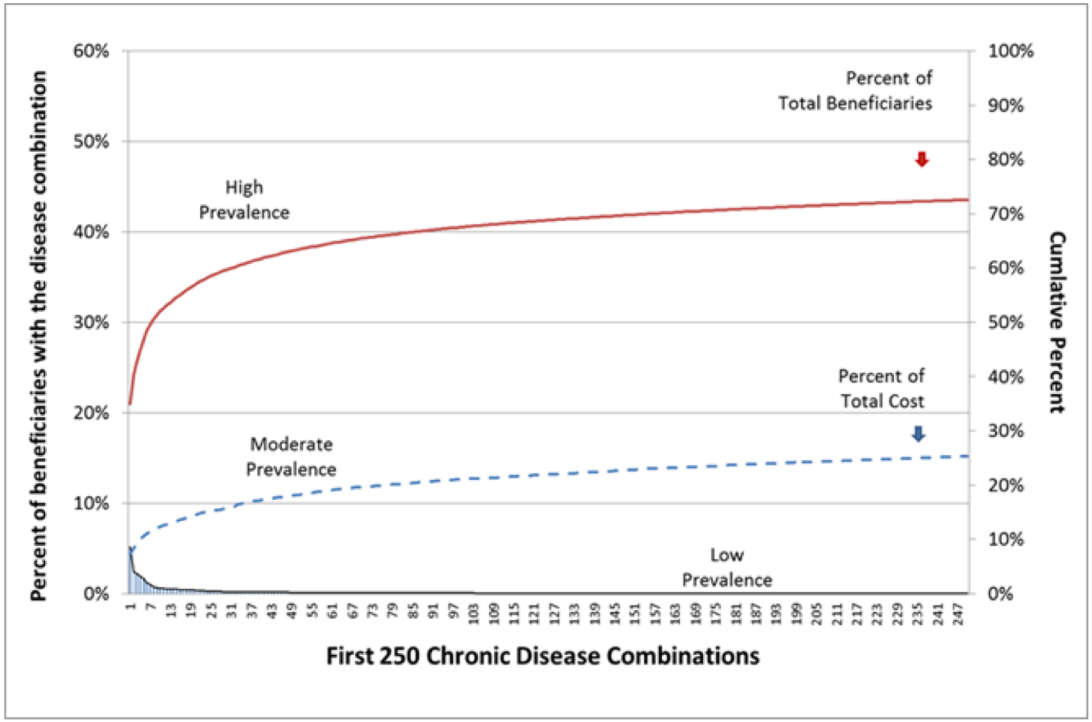
Small Patient Cell Sizes

- Given approximately 72,000 primary care providers*, then:
  - For the most prevalent DC (HCC 19) there are 23 beneficiaries per provider.
  - For the 100th most prevalent DC (HCCs 38 & 55) there are 0.26 beneficiaries per provider.

* Relationship of primary care physicians' patient caseload with measurement of quality and cost performance. Nyweide DJ, Weeks WB, Gottlieb DJ, Casalino LP, Fisher ES. *JAMA*. 2009.

Long Tailed Distribution of Medicare DCs

The graph displays the first 250 Diseases Combinations, ranked by prevalence, from the baseline HCC analysis. Note that the left Y-axis represents the proportion of the population that is included in each unique disease combination (black line). The right Y-axis represents the cumulative percent of the total population (red line) and the total expenditure (blue line), and is adjusted for the 32% of beneficiaries and 6% of expenditures that are associated with the no-HCC population. As there are over 2 million disease combinations calculated by this methodology, the figure's X-axis would need to be extended over 8,000 fold to the reader's right before both cumulative lines reached 100%.



Can We Prioritize Prevalent Conditions?

- **No!**
  - Restricting analysis to the 20 most prevalent of the 70 HCCs yields 53,476 DCs covering 40% of the population and 27% of expenditures.
  - Combined with the no-HCC group the 20 prevalent HCC DCs covers 75% of the population and 33% of expenditures.
  - Still missing 25% of the population and 67% of expenditures.
- Less common and rare diseases in aggregate are important drivers of expenditures.
- **This accounts for the limited performance of prevalence-based quality measures and current CMMI models**.

Rare Diseases

- In the UK it is estimated that 1 in 17 people have a rare disease, and rare diseases are a current focus of the UKs healthcare system.
  - Diagnosis is often long and costly.
  - Diseases are often scientifically informative.
  - The sharing of knowledge and the development of support groups difficult.

Twin Study

- Twins should have lower disease variation.
- In collaboration with VCU's Mid-Atlantic Twin Registry, we matched 396 pairs of MZ or "Identical" twins and 378 pairs of DZ or "Fraternal" twins to their Medicare claims data from 1991 through 2011.
- Studied pairs were predominantly white, male and Mid-Atlantic, and only included individuals in which both members survived to age 65.

Twin Study

- In a first of its kind study design, we used the Medicare claims database to construct unrelated demographically (sex, age, race and current county of residence) matched control pairs (MCPs) for both the MZ and DZ twins.
- We now have 4 groups to compare:
  - Monozygotic twins: MZ
  - Dizygotic twins: DZ
  - MZ Demographically Matched Control Pairs: MZ-MCP
  - DZ Demographically Matched Control Pairs: DZ-MCP

MCP Methodology Advantages

Group	Familial Genetics	Shared Family Environment	Controls For Demographics
MZ	100%	Yes	No
DZ	50%	Yes	No
MZ-MCP	0%	No	Yes
DZ-MCP	0%	No	Yes

Twin Study Results: Shared HCCs

- MZ (identical) twins shared 6.5% more HCCs than their MZ-MCP (26.3% vs. 19.8%, P<0.001).
- DZ (fraternal) twins shared 3.8% more HCCs than their DZ-MCP (25.6% vs. 21.8%, P<0.001).
- MZ-MCP/DZ-MCP (19.8% vs. 21.8%, p= 0.029)
- MZ/DZ (26.3% vs. 25.6%, p=0.52)

Twin Study Disease Correlation Summary

HCC (#)	Arrhythmias (92)	Stroke (96)	Diabetes & Renal (15)	Polyneuropathy (71)
MZ vs. MZ-MCP	+*	+*	+	+
MZ vs. DZ	-	-	+	+
DZ vs. DZ-MCP	+*	-	-	-

* Not Significant if ICD-9-CM code 427.3 (Atrial Fibrillation and Flutter) is excluded from the analysis.

Twin vs MCP: KS-Test of MED Curves

Abbreviations:

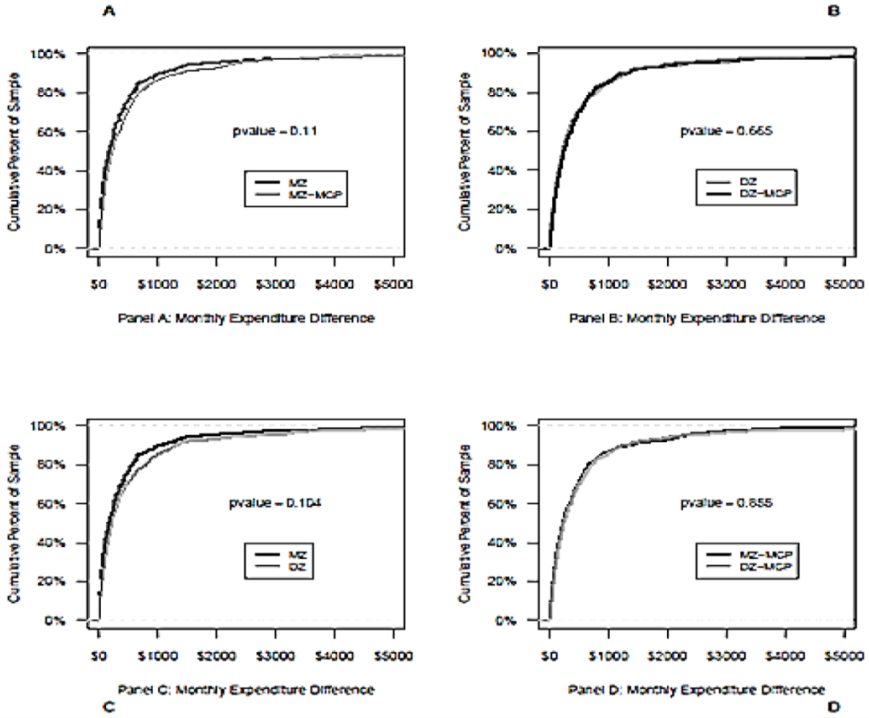
MZ: Monozygotic twin group

DZ: Dizygotic twin group

MZ-MCP: Monozygotic matched control pair group

DZ-MCP: Dizygotic matched control pair group

KS-Test: P-values were calculated using the Kolmogorov–Smirnov (KS) test



A Comparison of Disease Burden Between Twins and Control Pairs in Medicare: Quantification of Heredity's Role in Human Health.
Sorace J, Rogers M, Millman M, Rogers D, Price K, Queen S, Worrall C, Kelman J. *Population Health Management*. 2015 Feb 6. [Epub ahead of print]
PMID: 25658666

Twin Study Conclusions

- MCP methodology is viable and gave results that where distinct from and inclusive of those found with the tradition MZ vs. DZ design.
- The role of both heredity and shared family environment is limited in the study population.
- However, due to multiple comorbidities, heredity may still account for 1 major disease for every 2 to 3 people (crude estimate).
- Our findings are consistent with others:
  - The predictive capacity of personal genome sequencing. Roberts NJ, Vogelstein JT, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. Sci Transl Med. 2012 May 9;4(133):133ra58. doi: 10.1126/scitranslmed.3003380. Epub 2012 Apr 2. PMID: 22472521
  - Identically Different: https://www.youtube.com/watch?v=1W5SeBYERNI

Limitations

- May not apply to the non-Medicare population
- Use of HCCs as a disease aggregator.
- For the Twin Study there was:
  - Limited sex, race and geographic diversity.
  - Limited number of twins.
  - Both twins survived until 65.
- For the Disease Combination Study:
  - Only 1-Year Timeframes were used.
  - The order of diseases was not considered.

What Does it All Mean?

- **The problem is to complex for centralized top down solutions**.
  - Long tailed distribution that lacks useful means and measures of variance.
  - Distribution changes nationally over time.
  - No one provider or ACO has extensive experience with these patients.
- Must move toward a decentralized crowd sourced knowledge management solution.
  - Systems that are optimized for rare diseases are a useful “North Star” as rare patients are the new normal.
  - Roll for AI/ML with human supervision.
- Critical need for APMs that improve diagnostic accuracy and reward data/knowledge sharing.

Moving From HIT to HICT

- APMs must move from health information technology (HIT) to health information and communications technology (HICT).
- Using HICT diagnostic teams can support two vital communication loops.
- 1st Stage APM: "Inner Loop" communications provides care for a specific patient within the health care organization (the subject of active research).
- 2nd Stage APM: "Outer Loop" consist of **selected communications** at the national level to support the care of a defined patient cluster.
  - May be built off current efforts to support federated models of clinical research such as PCORI and OHDSI (note that the distinction between research and direct patient care will blur).
  - May incorporate distributed clinical crowd sourcing functions such as those being used in Project ECHO.

HICT

- **Correct diagnosis is a critical requirement for useful communications across the healthcare enterprise.**
- Follows treatment progress and outcomes locally (inner loop) while querying and comparing nationally (outer loop).
- **HICT is NOT useless emails and worthless reminders! These are examples of health information distraction technology!**

Summary of Possible APM Strategy

- 1st Stage Diagnostic APM might be a PTAC proposal based on improving diagnosis and communication within current ACOs such as:
  - Diagnostic Management Teams.
  - Greater collaboration between pathology and radiology in cancer diagnosis (consider greater collaboration more generally).
  - Confirm diagnosis for disease/condition associated APMs (e.g. chronic renal disease).
  - **Initially prioritize "inner loop communications."**

1st Stage APM

- Recent work has begun to address how a 1st generation APM might be implemented. Payment Innovations To Improve Diagnostic Accuracy And Reduce Diagnostic Error. Berenson R, Singh H. Health Aff (Millwood). 2018 Nov;37(11):1828-1835. doi: 10.1377/hlthaff.2018.0714.
  - Change Medicare fee schedule to include billing codes for improved communications as well as for diagnostic management teams.
  - Reduce documentation barriers and greater reward for cognitive work.
  - Make ACOs accountable for diagnostic timeliness and accuracy.
  - Condition based alternative payment models should assume the risk of correct diagnosis.

PTAC

- The Physician-Focused Payment Model Technical Advisory Committee (PTAC):
  - Created by MACRA in 2015 to encouraged the development of APMs referred to as physician-focused payment models (PFPMs).
  - PTAC's board consists of 11 members that meet publicly on a quarterly schedule.
- Secretary of HHS is required to respond to models forwarded by PTAC.
  - Models approved by the Secretary can be forwarded to CMMI for possible future development.

2nd Stage APM

- 2nd stage would be to develop “Fee For Knowledge Sharing” APMs for rare diseases and complex patients.
  - Would seek payments for coordinating diagnosis with both internal providers as well as collaboration with external providers with similar patient problems.
  - Would crowd source access to knowledgeable providers for consultation.
  - Prioritizes “outer loop communications.”
- ACOs would have an allowance to pay for knowledge and data from outside organizations.
  - The ACO cannot keep it if it does not spend it.
  - The ACO can only receive a payment when it shares knowledge requested by others.
  - Queries must originate from a specific patient(s).

2nd Stage APM

- Hybrid payment models: Capitated (e.g. Medicare Part-C) payment for direct patient care coupled with a Fee for Knowledge Sharing Component.
- Develop CMMI proposal. Consider initially enrolling Medicare Part-C plans especially if they have limited geographic overlap.

Thank You!

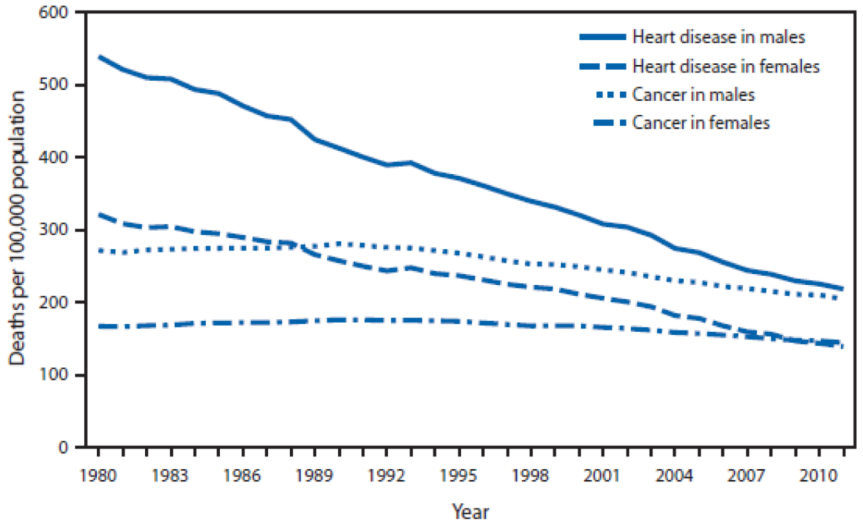
James Sorace MD MS
jamessorace1@gmail.com

Heredity vs. General Environment

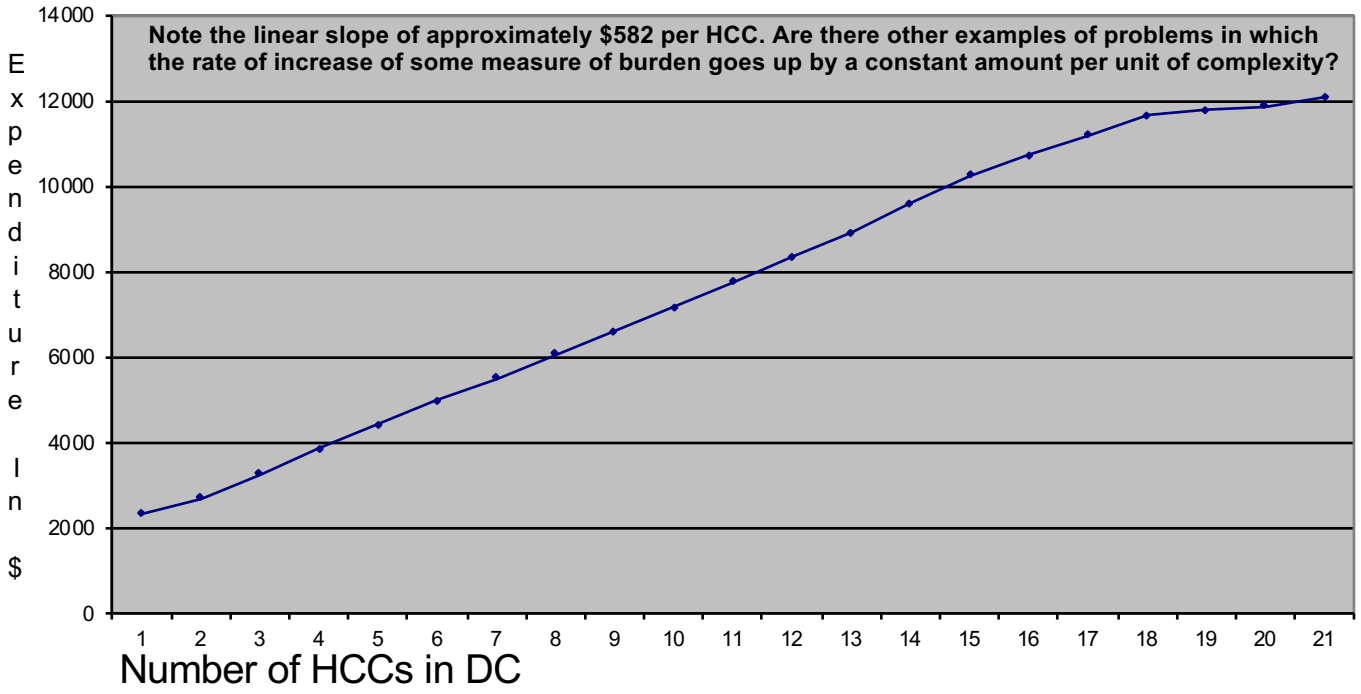
The twin study overlaps a period of time in which there was a significant reduction in deaths due to vascular disease.

Widespread adoption of Statins, and decreased smoking rates, among other public health measures, also occurred in this period.

https://www.cdc.gov/mmwr/preview/mmwrhtml/figures/m6337qsf.gif



Expenditure per Individual per HCC



Do Disease Combinations Change Over Time?

Stable DCs are the set of DCs that were present across 3 consecutive years of analysis (2007 to 2009).

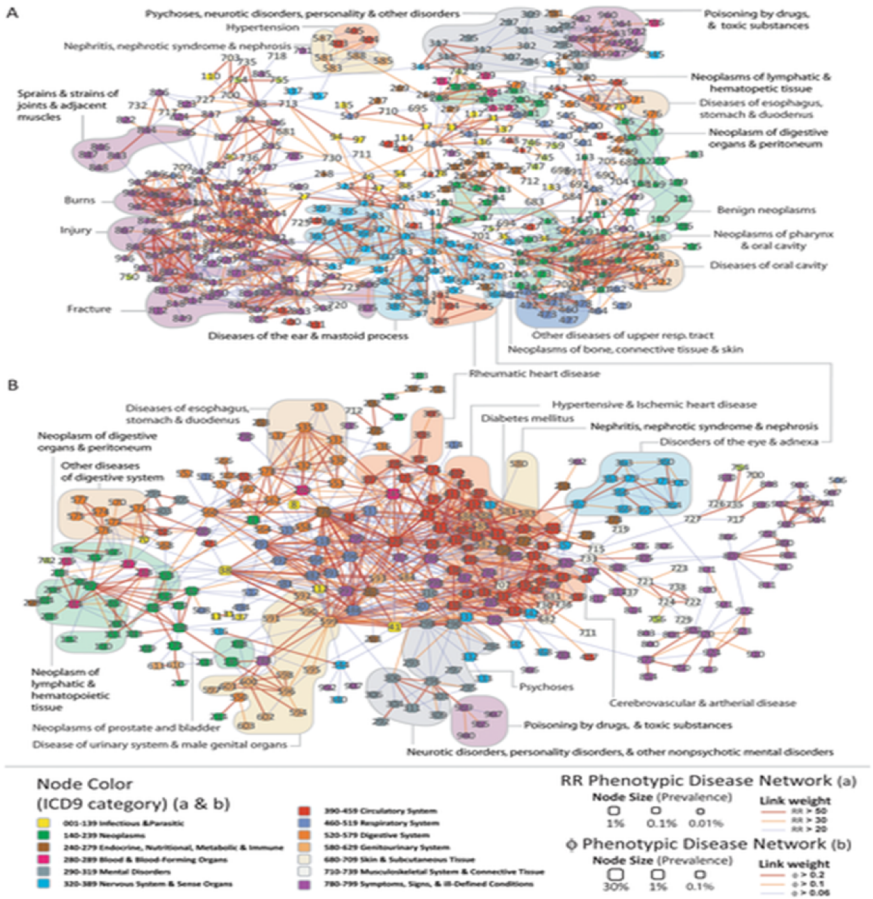
Note that the percent of stable disease combinations exceeds the percent of occupied DCs.

Might stable DCs may be enriched for gene/environment interactions?

Number of HCCs In DC	Number of Mathematical Combinations of 70	Number of Observed DCs	% of Occupied DCs	% of Stable DCs
1	70	70	100%	100%
2	2,415	2,309	96%	98%
3	54,740	31,475	57%	82%
4	916,895	143,651	16%	57%
5	12,103,014	288,137	2%	36%
6	131,115,985	353,835	0%	22%
7	1,198,774,720	335,491	0%	14%

Phenotypic Disease Networks (PDNs)

Nodes are diseases; links are correlations. Node color identifies the ICD9 category; node size is proportional to disease prevalence. Link color indicates correlation strength. Figure A. PDN constructed using *RR*. Only statistically significant links with RR_{ij}>20 are shown. Figure B. PDN built using φ-correlation. Here all statistically significant links where φ>0.06 are shown.



Hidalgo CA, Blumm N, Barabási A-L, Christakis NA (2009) A Dynamic Network Approach for the Study of Human Phenotypes. *PLoS Computational Biology* 5(4): e1000353. doi:10.1371/journal.pcbi.1000353

http://www.ploscompbiol.org/article/info:doi/10.1371/journal.pcbi.1000353