DISEASE

DEFINITION:

ALTERATION (REDUCTION, INCREASE, LACK) OF CELLULAR FUNCTION OF CELLS/TISSUES/ORGANS

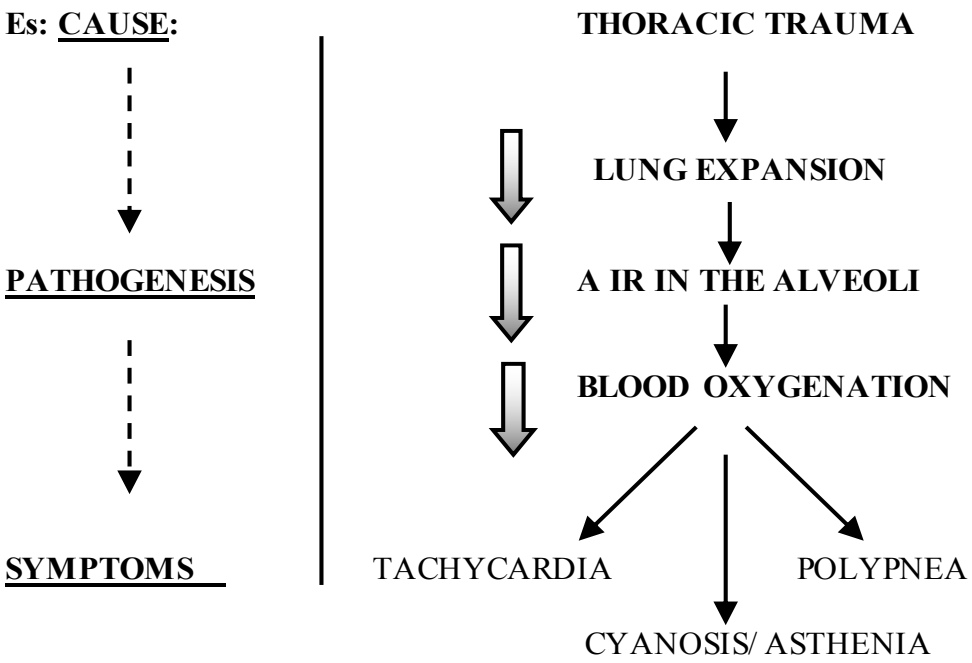
ALTERATION OF HOMEOSTATIC EQUILIBRIUM

PATHOLOGY

THE STUDY OF DISEASE

ETIOLOGY

PATHOGENESIS



ETIOLOGY: CAUSES OF DISEASE

CONGENITAL: START BEFORE OR IN CONCOMITANCE WITH BIRTH; can be:

- **Genetical,**
- **Due to pregnancy**
- **Due to delivery**

HEREDITARY

ACQUIRED (after birth)

- CHEMICAL

- Exogenous compounds (either natural or deriving from human activities)
- Endogenous molecules
  - -Metabolic/catabolic products (bilirubin, lactic acid)
  - -Reactive oxygen and nitrogen species (ROS, RNS)
  - -Modified molecules (oxidized lipoproteins, glycated proteins)

- PHYSICAL

- temperature
- radiations
- pressure
- noise
- electrical

- BIOLOGICAL

Direct or indirect damage deriving from pathogen overgrowth

- DEFICIENCIES OR EXCESSES

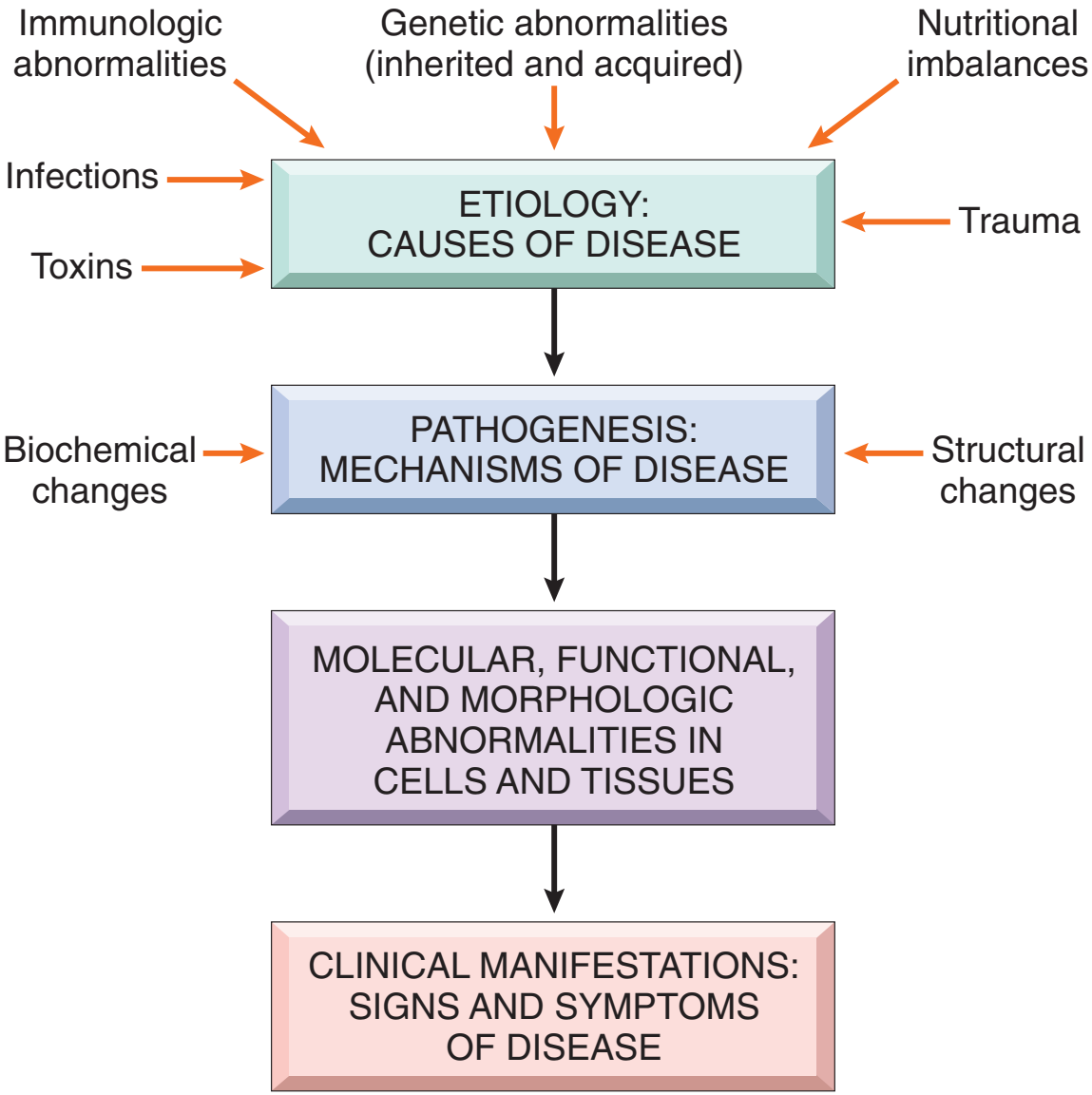
- oxygen deficiency
- vitamin deficiency
- iron deficienzy/overload
- nutriens deficiency/excess
- cholesterol accumulation

- IMMUNOLOGICAL

- IATROGENIC

- drugs

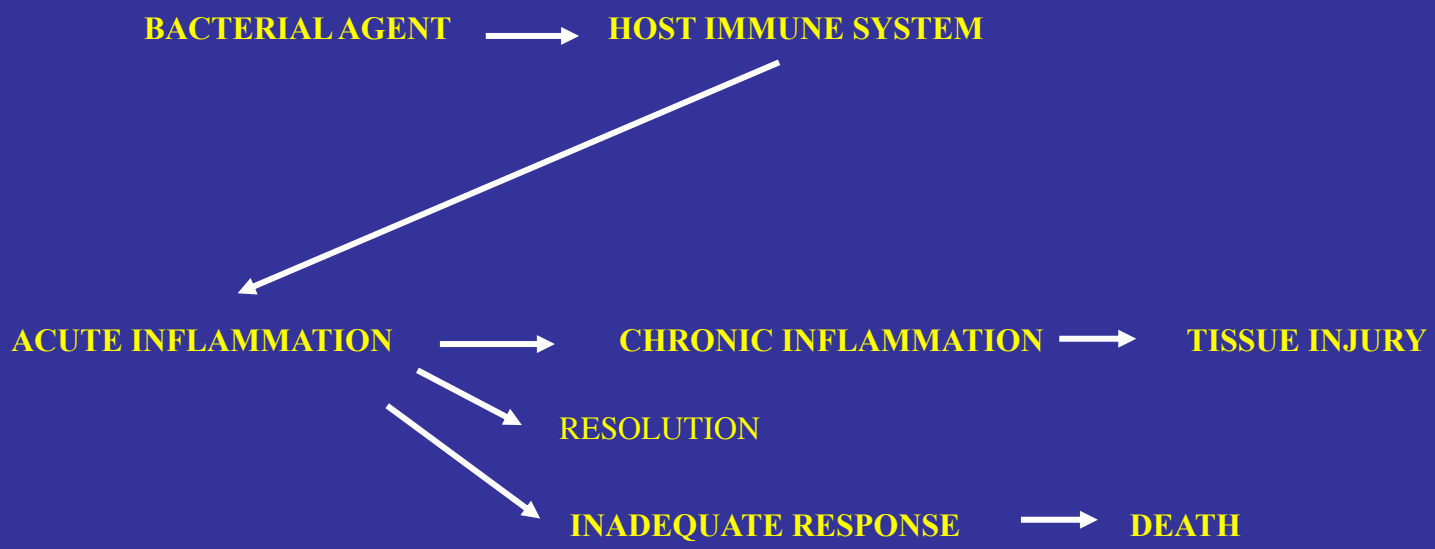
Steps in the evolution of a disease



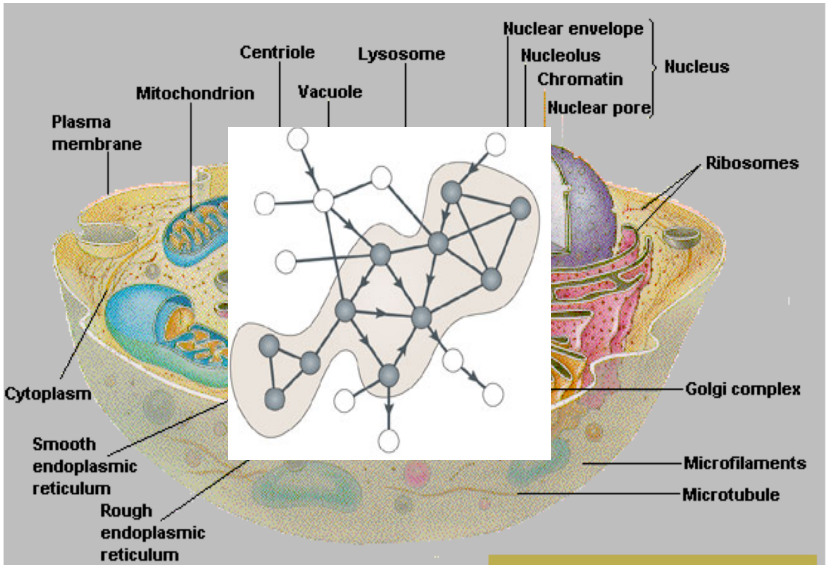
PATOGENESIS

Mechanisms of development and progression of disease, which account for the cellular and molecular changes that give rise to the specific functional and structural abnormalities that characterize any particular disease.

ES:



Do the interactions between cell molecules generate comorbidity patterns at population level?



Network medicine

Applications:

The impact of cellular networks on comorbidity

The impact of cellular networks on comorbidity

Molecular Systems Biology 5; Article number 262; doi:10.1038/msb.2009.16
Citation: ***Molecular Systems Biology*** **5:262**

www.molecularsystemsbiology.com

molecular systems biology

REPORT

The impact of cellular networks on disease comorbidity

Juyong Park[1,2,*], Deok-Sun Lee[1,2,4], Nicholas A Christakis[3] and Albert-László Barabási[1,2,5,*]

[1] Departments of Physics, Biology, and Computer Science, Center for Complex Network Research, Northeastern University, Boston, MA, USA, [2] Center for Cancer Systems Biology, Dana-Farber Cancer Institute, Boston, MA, USA, [3] Department of Health Care Policy, Harvard Medical School, Boston, MA, USA, [4] Department of Natural Medicine Sciences, Inha University, Incheon, Korea and [5] Department of Medicine, Harvard Medical School, Boston, MA, USA
* Corresponding authors. J Park or L Barabási, Departments of Physics, Biology, and Computer Science, Center for Complex Network Research, Northeastern University 360 Huntington Ave, Boston, MA 02115, USA. Tel.: +1 617 373 7774; Fax: +1 617 373 4385; E-mails: perturbation@gmail.com or barabasi@gmail.com

Received 24.7.08; accepted 25.2.09

The impact of disease-causing defects is often not limited to the products of a mutated gene but, thanks to interactions between the molecular components, may also affect other cellular functions, resulting in potential comorbidity effects. By combining information on cellular interactions, disease–gene associations, and population-level disease patterns extracted from Medicare data, we find statistically significant correlations between the underlying structure of cellular networks and disease comorbidity patterns in the human population. Our results indicate that such a combination of population-level data and cellular network information could help build novel hypotheses about disease mechanisms.
Molecular Systems Biology 7 April 2009; doi:10.1038/msb.2009.16
Subject Categories: bioinformatics; molecular biology of disease
Keywords: cellular networks; comorbidity; database; population-level statistics

Morbidity (from Latin *morbidus*, meaning "sick, unhealthy")
= is a diseased state, disability, or poor health due to any cause.
= incidence of a particular disease in a population

Morbilità/Morbosità indica uno stato patologico dell'individuo

Comorbidity = (comorbosità in italiano) coexistence of more pathological states/diseases in the same person

Comorbidity – general considerations and open questions

1. Cellular functions are controlled by networks of genes, proteins, metabolites, etc. Defects in one module can “propagate” to other modules (everything is interconnected!)

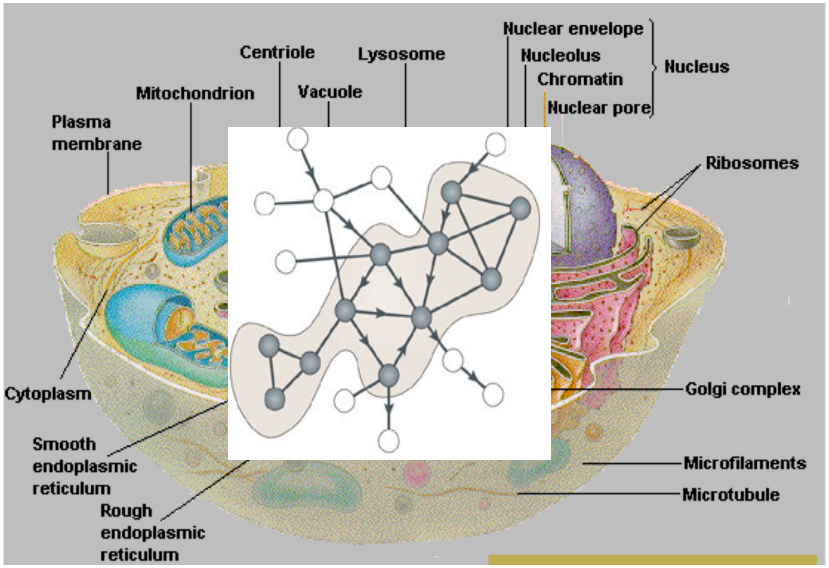
2. Are the interdependencies between cellular/subcellular networks important only for the individual or also for the global population?

3. Do the diseases linked in the human disease network present comorbidity?

4. Do the diseases that share genes and proteins that interact present comorbidity?

5. Do the diseases that have high number of co-expressed genes show comorbidity?

To what degree the topological connectivity of cellular networks is related to the manifestation of human diseases, possibly leading to phenotypic interdependencies and comorbidity?

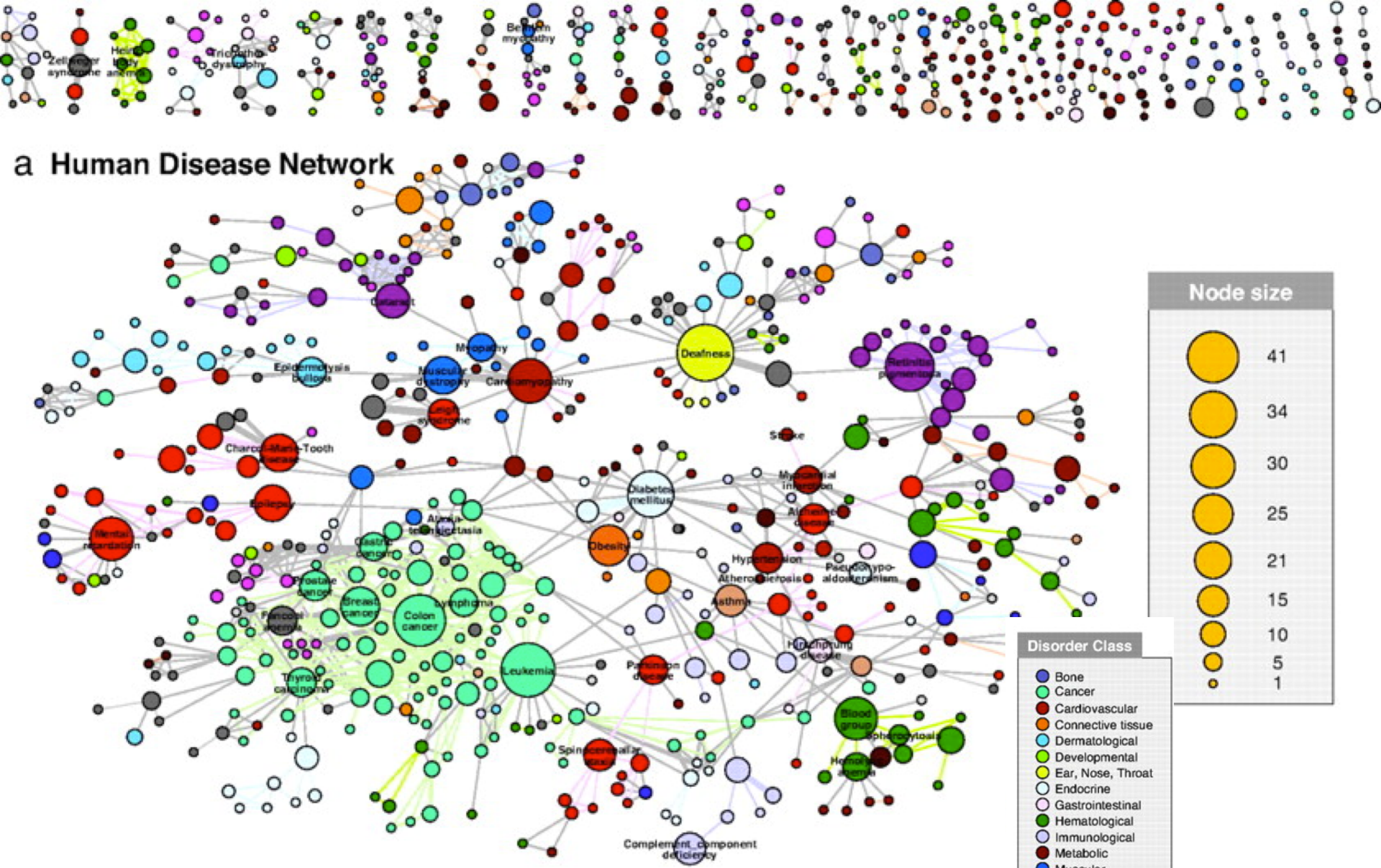


Steps of the comorbidity study of “genetic” diseases

1. Building the genetic disease network

2. Analysis of comorbidity

3. Predictions / Conclusions



1. Building the genetic disease network

Goh K et al. PNAS 2007;104:8685-8690

The properties of disease network (HDN)

1. Diseases and disease classes are strongly

interconnected) (from 1284 diseases, 867 have at least one link)

2. 516 diseases form a giant cluster, suggesting shared genetical bases for numerous diseases

2. Analysis of comorbidity

Disease classification

"The International Classification of Diseases, 9th Revision, Clinical Modification" (ICD-9-CM), issued for use beginning October 1, 2008 for federal fiscal year 2009 (FY09). http://www.cms.gov/
http://www.cms.gov/MedicareGenInfo/ http://www.medicare.gov/

The ICD-9-CM is maintained jointly by the National Center for Health Statistics (NCHS) and the Centers for Medicare & Medicaid Services (CMS).

http://www.icd9data.com/

001-139 INFECTIOUS AND PARASITIC DISEASES
140-239 NEOPLASMS
240-279 ENDOCRINE, NUTRITIONAL AND METABOLIC DISEASES, AND IMMUNITY DISORDERS
280-289 DISEASES OF THE BLOOD AND BLOOD-FORMING ORGANS
290-319 MENTAL DISORDERS
320-389 DISEASES OF THE NERVOUS SYSTEM AND SENSE ORGANS
390-459 DISEASES OF THE CIRCULATORY SYSTEM
460-519 DISEASES OF THE RESPIRATORY SYSTEM
520-579 DISEASES OF THE DIGESTIVE SYSTEM
580-629 DISEASES OF THE GENITOURINARY SYSTEM
630-679 COMPLICATIONS OF PREGNANCY, CHILDBIRTH, AND THE PUERPERIUM
680-709 DISEASES OF THE SKIN AND SUBCUTANEOUS TISSUE
710-739 DISEASES OF THE MUSCULOSKELETAL SYSTEM AND CONNECTIVE TISSUE
740-759 CONGENTIAL ANOMALIES
760-779 CERTAIN CONDITIONS ORIGINATING IN THE PERINATAL PERIOD
780-799 SYMPTOMS, SIGNS, AND ILL-DEFINED CONDITIONS
800-999 INJURY AND POISONINGRY AND POISONING

2011 ICD-9-CM Diagnosis Code 170
Malignant neoplasm of bone and articular cartilage

2011 ICD-9-CM Diagnosis Code 170.0
Malignant neoplasm of bones of skull and face except mandible

2011 ICD-9-CM Diagnosis Code 170.1
Malignant neoplasm of mandible

2011 ICD-9-CM Diagnosis Code 170.2
Malignant neoplasm of vertebral column excluding sacrum and coccyx

2011 ICD-9-CM Diagnosis Code 170.3
Malignant neoplasm of ribs sternum and clavicle

2011 ICD-9-CM Diagnosis Code 170.4
Malignant neoplasm of scapula and long bones of upper limb

2011 ICD-9-CM Diagnosis Code 170.5
Malignant neoplasm of short bones of upper limb

2011 ICD-9-CM Diagnosis Code 170.6
Malignant neoplasm of pelvic bones sacrum and coccyx

2011 ICD-9-CM Diagnosis Code 170.7
Malignant neoplasm of long bones of lower limb

2011 ICD-9-CM Diagnosis Code 170.8
Malignant neoplasm of short bones of lower limb

2011 ICD-9-CM Diagnosis Code 170.9
Malignant neoplasm of bone and articular cartilage site unspecified

2. Analysis of comorbidity

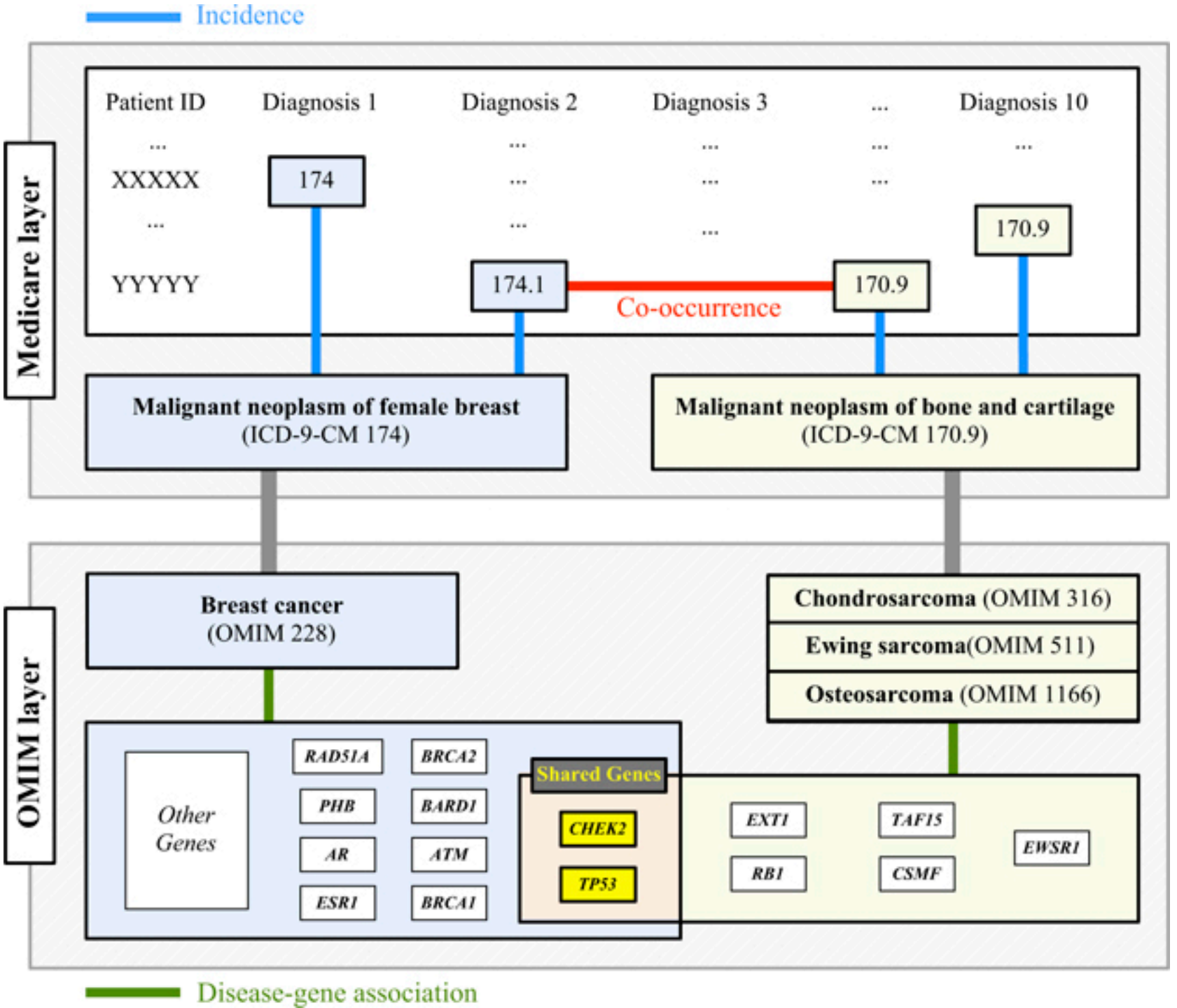
Patients

Study made on patients registered in Medicare database :

N= 13.039.018 pazienti registrati su Medicare per un totale di 32.341.348 visite mediche
Periodo: più di 4 anni (1990 to 1993)

Età media: 76.3 ± 7.4;
41.8% maschi; 89.9% caucasici

http://www.medicare.gov/

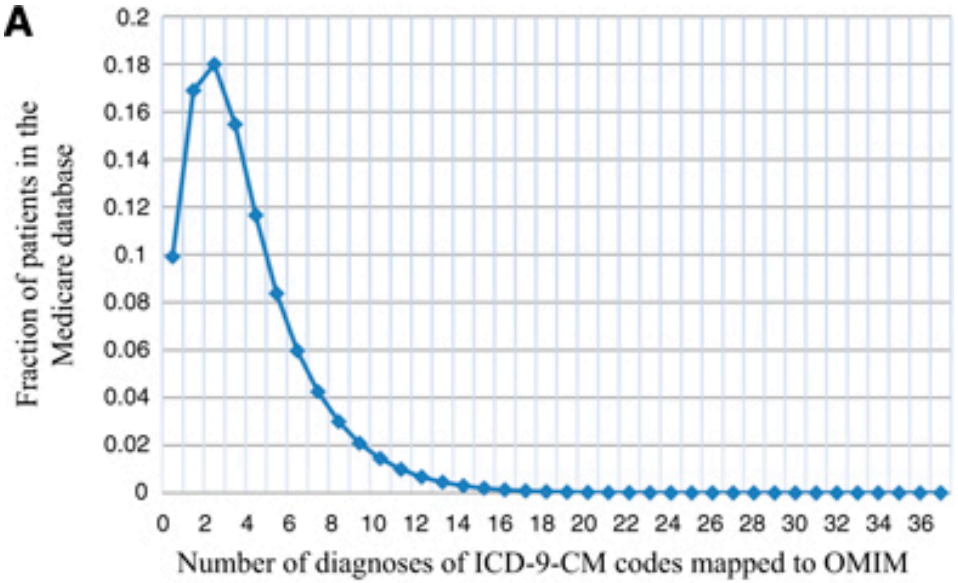

Incidence
Medicare layer
Patient ID
Diagnosis 1
Diagnosis 2
Diagnosis 3
...
Diagnosis 10
XXXXX
YYYYY
174
174.1
170.9
170.9
Co-occurrence
Malignant neoplasm of female breast
(ICD-9-CM 174)
Malignant neoplasm of bone and cartilage
(ICD-9-CM 170.9)
OMIM layer
Breast cancer
(OMIM 228)
Chondrosarcoma (OMIM 316)
Ewing sarcoma(OMIM 511)
Osteosarcoma (OMIM 1166)
Other Genes
RAD51A
BRCA2
PHB
BARD1
AR
ATM
ESR1
BRCA1
Shared Genes
CHEK2
TP53
EXT1
TAF15
EWSR1
RB1
CSMF
Disease-gene association

>4.900 disease-gene associations on OMIM database (2008)

ICD-9-CM codes (12.000) were manually connected
To the name of diseases in OMIM database

Only 763 ICD-9-CM codes were found in OMIM database!

90% of Medicare patients
Were diasgnosed with
ICD – OMIM codes (763)!!



COMORBIDITY CALCULATION

i, j = co-expressed diseases

I_i = incidence of disease i (nr. of patients)

C_{ij} = number of patients simultaneously diagnosed with diseases i and j

Parameters to measure comorbidity:

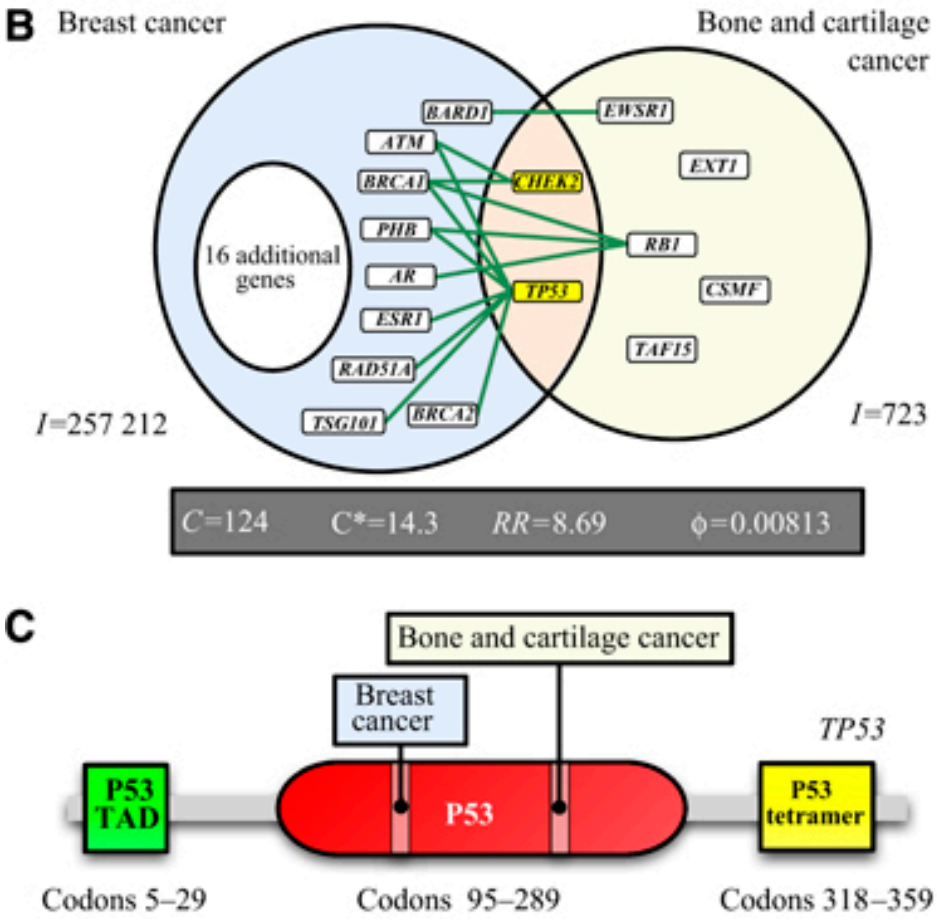
***RR* = C_{ij}/C_{ij}^* = *relative risk* (rischio relativo)** $RR>1$

C_{ij}^* = I_iI_j/N (expected value of C_{ij} when the two diseases are independent)

ϕ correlation $\phi>0$

$$\phi_{ij} = \frac{NC_{ij} - I_iI_j}{\sqrt{I_iI_j(N-I_i)(N-I_j)}},$$

Example of two diseases that show comorbidity and are linked at the level of cellular network



The two diseases share 13 PPI of the gene products

The two diseases show domain-sharing In TP53 protein

Results

From the analysis of ICD and OMIM databases and PPI networks, in the application (2008) were identified 83.924 pairs of co-expressed diseses.

2.239 pairs were linked by genes or PPIs

Parameters for the quantification of the relationship between cellular networks and comorbidity

(n_{ij}^{g}) **Number of disease genes** shared between diseases *i* and *j*; Quantifies the potential common genetic origin of the two diseases $n_{ij}^{g} = |G_i \cap G_j|$.

(n_{ij}^{p}) **Number of PPIs** between the proteins (gene products) of diseases i and j; Captures the PPI network-level relationships between them the proteins encoded by disease genes (non only by shared genes)

$(\bar{\rho}_{ij})$ **Average Pearson correlation of coexpression** between pairs of genes from each disease; Captures the degree to which the genes associated with the two diseases are coexpressed (see paper by Ge at al., *Genomics* 2005)

$$\rho_{ab} = \frac{n_t \sum_t x_{at} x_{bt} - \sum_t x_{at} \sum_t x_{bt}}{\sqrt{(n_t \sum_t x_{at}^2 - [\sum_t x_{at}]^2)(n_t \sum_t x_{bt}^2 - [\sum_t x_{bt}]^2)}}.$$

Useful published study to determine gene coexpression

Genomics. 2005 Aug;86(2):127-41.

Interpreting expression profiles of cancers by genome-wide survey of breadth of expression in normal tissues.

Ge X, Yamamoto S, Tsutsumi S, Midorikawa Y, Ihara S, Wang SM, Aburatani H.

Abstract

A critical and difficult part of studying cancer with DNA microarrays is data interpretation. Besides the need for data analysis algorithms, integration of additional information about genes might be useful. **We performed genome-wide expression profiling of 36 types of normal human tissues** and identified 2503 tissue-specific genes. We then systematically studied the expression of these genes in cancers by reanalyzing a large collection of published DNA microarray datasets. We observed that the expression level of liver-specific genes in hepatocellular carcinoma (HCC) correlates with the clinically defined degree of tumor differentiation. Through unsupervised clustering of tissue-specific genes differentially expressed in tumors, we extracted expression patterns that are characteristic of individual cell types, uncovering differences in cell lineage among tumor subtypes. We were able to detect the expression signature of hepatocytes in HCC, neuron cells in medulloblastoma, glia cells in glioma, basal and luminal epithelial cells in breast tumors, and various cell types in lung cancer samples. We also demonstrated that tissue-specific expression signatures are useful in locating the origin of metastatic tumors. Our study shows that integration of each gene's breadth of expression (BOE) in normal tissues is important for biological interpretation of the expression profiles of cancers in terms of tumor differentiation, cell lineage, and metastasis.

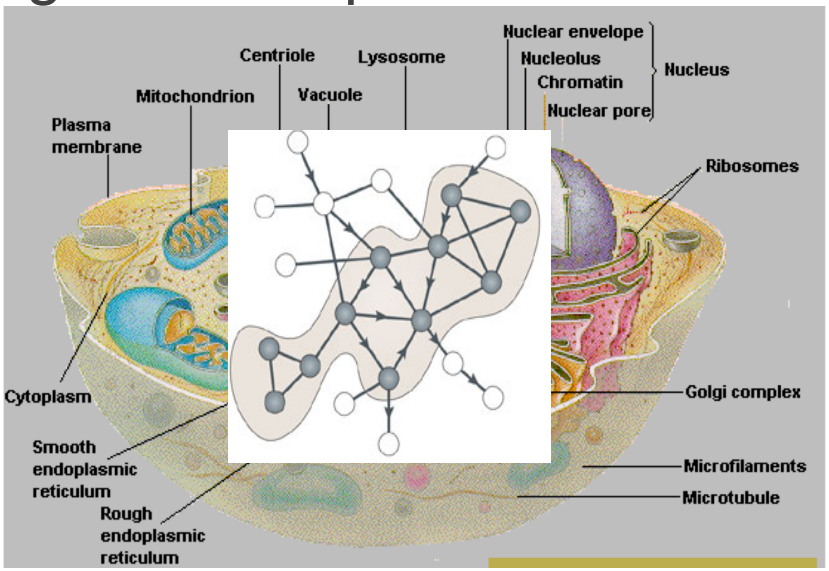
How to measure Gene Co-expression

- **Pearson's correlation coefficient (most widely used)**,
- - Mutual Information,
- - Spearman's rank correlation coefficient
- Euclidean distance

The PCC takes a value between -1 and 1 (values close to 1 show strong correlation).
Positive values correspond to an activation mechanism where the expression of one gene increases with the increase in the expression of its co-expressed gene and vice versa.
Negative values: when the expression value of one gene decreases with the increase in the expression of its co-expressed gene (underlying suppression mechanism?)

To what degree the topological connectivity of cellular networks is related to the manifestation of human diseases, possibly leading to phenotypic interdependencies and comorbidity?

Nr. of shared genes, PPIs, gene co-expression



Comorbidity

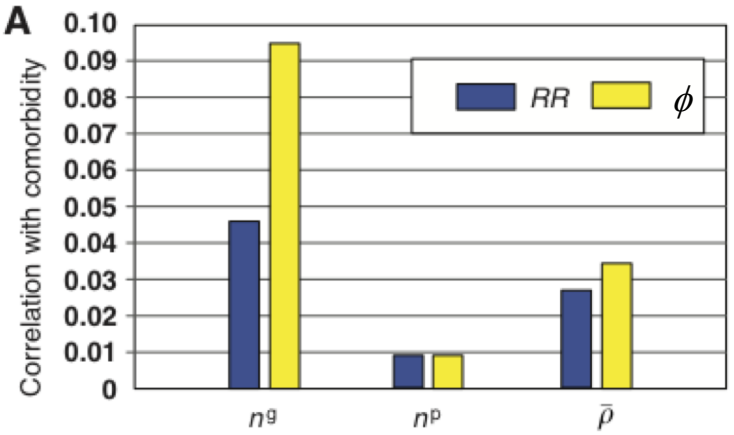
Cellular networks correlate with comorbidity

Pearson correlation between comorbidity for all disease pairs and n_g, n_p e $\overline{\rho}$

Table I The Pearson correlation between relative risk, ϕ-correlation and the three genetic variables (see Figure 2A)

Genetic variables	Pearson correlation with relative risk	Pearson correlation with ϕ-correlation
n^g	0.0469 ($P \approx 3.85 \times 10^{-4}$)	0.0902 ($P \approx 1.48 \times 10^{-4}$)
n^p	0.00948 ($P \approx 1.65 \times 10^{-2}$)	0.00941 ($P \approx 1.49 \times 10^{-2}$)
$\bar{\rho}$	0.0272 ($P \approx 1.07 \times 10^{-3}$)	0.0334 ($P \approx 3.41 \times 10^{-4}$)

All correlations are positive, with *P*-values shown in parentheses.



Positive correlation (positive PCC) with all three parameters.
n^g shows the highest correlation.

However, **the correlation is weak** because from 83.924 disease pairs found on ICD and consideredfor this study, only 2.239 disease pairs share genes or proteins

The correlation obtained between cellular networks and comorbidity is low

Limits of the application:

The values obtained do not consider:
- **The environment**
- **Life stile**
- **Various treatments**

We do not know all the associations between genes and diseases from OMIM database

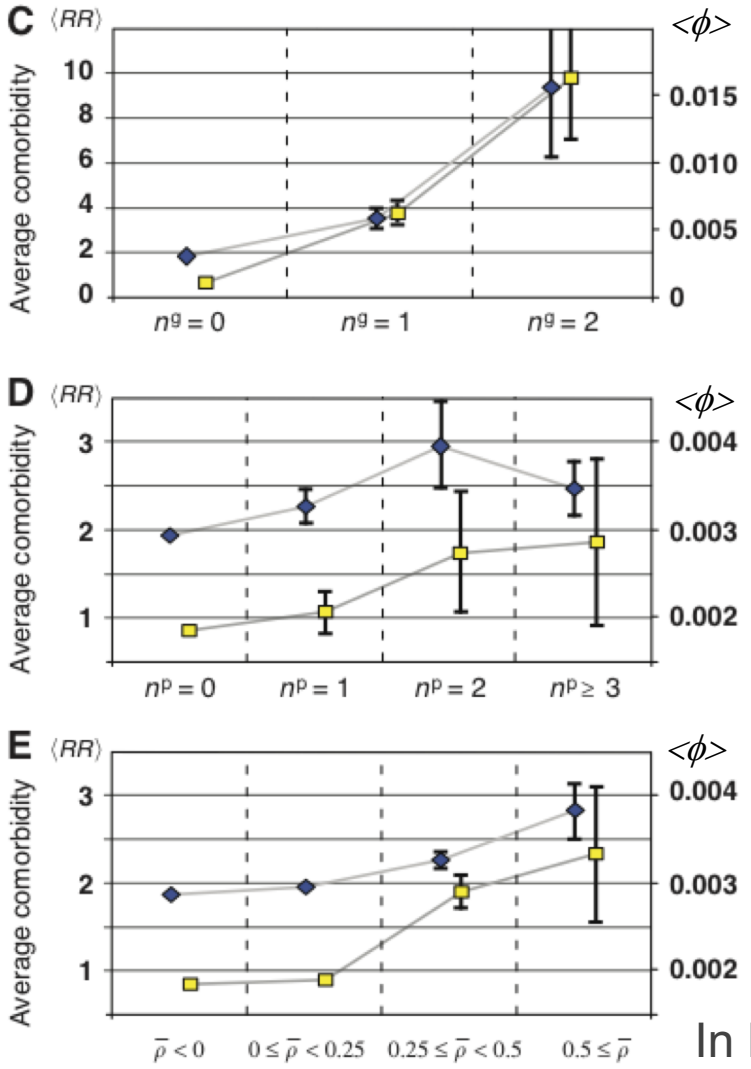
There is “noise” between OMIM and ICD-9-CM codes

Shared disease genes and proteins are known only for a minority of diseases showing comorbidity

Cellular networks correlate with comorbidity

Mean comorbidity calculated for disease pairs with increassing value of n^g, n^p, $\bar{\rho}$

The diseases with more connections have a higher comorbidity



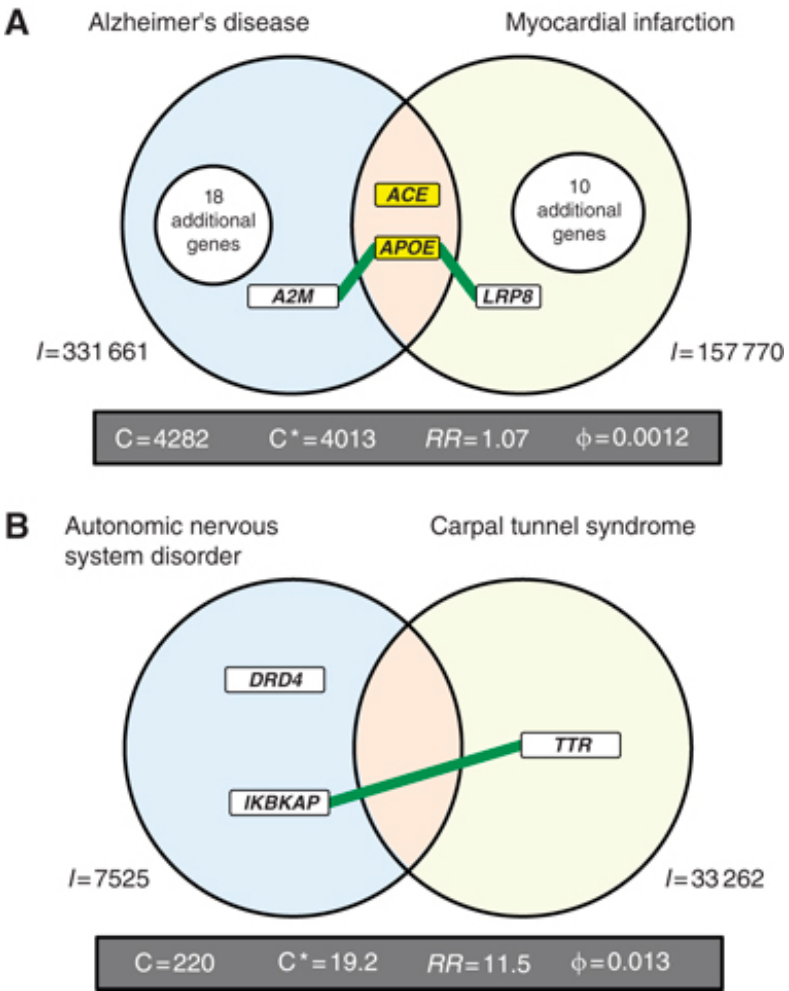
In blue: RR
In yellow: ϕ

3. Predictions / Conclusions

Cellular networks correlate comorbidity and predict:

- Known Comorbidities: a) diabetes-obesity; b) breast cancer-osteosarcoma
- New Comorbidities: a) Alzheimer's disease and myocardial infarction; b) Autonomic nervous system disorder and carpal tunel syndrome

***Gene* – proteina codificata**
ACE-angiotensin-converting enzyme
APOE- apolipoproteina E
TTR-transthyretin
IKBKAP- IKAP protein
A2M – alpha-2 macroglobulin
LRP8 – low density lipoprotein receptor-related protein8



UniProtKB/Swiss-Prot: A2MG_HUMAN, P01023
Function: Is able to inhibit all four classes of proteinases by a unique 'trapping' mechanism. This protein has a peptide stretch, called the 'bait region' which contains specific cleavage sites for different proteinases. When a proteinase cleaves the bait region, a conformational change is induced in the protein which traps the proteinase. The entrapped enzyme remains active against low molecular weight substrates (activity against high molecular weight
substrates is greatly reduced). Following cleavage in the bait region a thioester bond is hydrolyzed and mediates the
covalent binding of the protein to the proteinase

Entrez Gene summary for LRP8:
This gene encodes a member of the low density lipoprotein receptor (LDLR) family. Low density lipoprotein receptors are
cell surface proteins that play roles in both signal transduction and receptor-mediated endocytosis of specific ligands for lysosomal degradation. The encoded protein plays a critical role in the migration of neurons during development by mediating Reelin signaling, and also functions as a receptor for the cholesterol transport protein apolipoprotein E. Expression of this gene may be a marker for major depressive disorder.

Apolipoprotein E (ApoE) is a class of apolipoprotein found in the chylomicron and Intermediate-density lipoprotein (IDLs) that is essential for the normal catabolism of triglyceride-rich lipoprotein constituents. In peripheral tissues, ApoE is primarily produced by the liver and macrophages, and mediates cholesterol metabolism in an isoform-dependent manner. In the central nervous system, ApoE is mainly produced by astrocytes, and transports cholesterol to neurons via ApoE receptors, which are members of the low density lipoprotein receptor gene family. APOE [2] is 299amino acids long and transports lipoproteins, fat-soluble vitamins, and cholesterol into the lymph system and then into the blood. It is synthesized principally in the liver, but has also been found in other tissues such as the brain, kidneys, and spleen.
APOE was initially recognized for its importance in lipoprotein metabolism and cardiovascular disease. Defects in APOE result in familial dysbetalipoproteinemia aka type III hyperlipoproteinemia (HLP III), in which increased plasma cholesterol and triglycerides are the consequence of impaired clearance of chylomicron, VLDL and LDL remnants[*citation needed*]. *More recently, it has been studied for its role in several biological processes not directly* ***related to lipoprotein transport, including Alzheimer's disease (AD), immunoregulation, and cognition.***

Entrez Gene summary for TTR**:**This gene encodes transthyretin, **one of the three prealbumins including alpha-1-antitrypsin, transthyretin andorosomucoid**. **TRANSTHYRETIN IS A CARRIER PROTEIN; IT TRANSPORTS THYROID HORMONES IN THE PLASMA AND CEREBROSPINAL FLUID,AND ALSO TRANSPORTS RETINOL (VITAMIN A) IN THE PLASMA**. The protein consists of a tetramer of identical subunits. Morethan 80 different mutations in this gene have been reported; **MOST MUTATIONS ARE RELATED TO AMYLOID DEPOSITION, AFFECTING PREDOMINANTLY PERIPHERAL NERVE AND/OR THE HEART, AND A SMALL PORTION OF THE GENE MUTATIONS IS NON-AMYLOIDOGENIC. THE DISEASES CAUSED BY MUTATIONS INCLUDE AMYLOIDOTIC POLYNEUROPATHY, EUTHYROID HYPERTHYROXINAEMIA,AMYLOIDOTIC VITREOUS OPACITIES, CARDIOMYOPATHY, OCULOLEPTOMENINGEAL AMYLOIDOSIS, MENINGOCEREBROVASCULAR AMYLOIDOSIS, CARPAL TUNNEL SYNDROME**.

IKBKAP (inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein) is a human gene that provides instructions to make the IKAP protein, which is found in a variety of cells throughout the body, including brain cells. Although the exact function of the IKAP protein is not clearly understood, it probably plays a role in transcription, which is the process of making a blueprint of a gene for protein production. Researchers have identified the IKAP protein as part of a six-protein complex (called the holo-elongator complex) that interacts with enzymes necessary for transcription. The IKAP protein probably performs other functions in the cell as well, such as responding to stress. Its homolog in fly (D-elp1) has RNA-dependent RNA polymerase activity and is involved in RNA interference.[1]
Related conditions

FAMILIAL DYSAUTONOMIA is caused by mutations in the IKBKAP gene. Nearly all individuals with familial dysautonomia have two copies of the same mutation in each cell, which causes part of the IKBKAP gene to be skipped during transcription. (This alteration is often called exon skipping.) This skipping mutation results in a decreased amount of IKAP protein in their cells. This mutation, however, behaves inconsistently. As a result, some cells produce near normal amounts of IKAP protein, and other cells (particularly cells in the nervous system) have very little IKAP protein.

Network medicine

Applicazioni:

L’impatto dei network cellulari sulla comorbidity