

Redes de colaboración científica en las publicaciones cardiológicas argentinas

RAÚL A. BORRACCI^{MTSAC, 1, 2}, HERNÁN C. DOVAL^{MTSAC, 1}, DIEGO MANENTE^{1, 1}, CARLOS D. TAJER^{MTSAC, 1}

Recibido: 25/08/2009

Aceptado: 07/10/2009

Dirección para separatas:

Dr. Raúl A. Borracci
La Pampa 3030 - 1° B
(1428) Buenos Aires, Argentina
e-mail:
borracci@universia.com.ar

RESUMEN

Introducción

La estructura de las redes de colaboración científica se ha estudiado recientemente en distintas disciplinas. La construcción de gráficos de colaboración entre investigadores se ha llevado a cabo para varias disciplinas a través del análisis de grandes bases de datos y mediante el auxilio de *software* especializado capaz de interconectar a los distintos autores y coautores en una gran red de cooperación científica. El uso de estas redes no es nuevo en el campo de la bibliometría; sin embargo, últimamente se ha prestado más atención a las redes de coautoría dentro de una comunidad académica, ya que éstas podrían ser más representativas de la estructura del conocimiento de dicha comunidad.

Objetivo

Describir la estructura de un conjunto de redes de colaboración científica en la Argentina a partir del análisis de coautoría de los artículos cardiológicos publicados durante 2007.

Material y métodos

Se realizó una búsqueda bibliográfica de artículos cardiológicos argentinos, que incluyó los aparecidos en *Medline* y en dos revistas locales: *Rev Argent Cardiol* y *Rev Fed Arg Cardiol*. Se construyeron las redes de colaboración entre autores por medio del algoritmo Kamada-Kawai en el *software* Pajek.

Resultados

El promedio de artículos por autor se halló entre 1,12 y 1,24, el exponente *tau* de productividad entre 2,78 y 3,45, el promedio de autores por artículo entre 3,60 y 6,51 y el de colaboradores por autor entre 2,60 y 4,88. La construcción de sendas redes de colaboración mostró que los componentes gigantes oscilaron entre el 13,1% y el 65,8%, la distancia promedio entre autores fue de 1,5 a 8,5 y la distancia máxima estuvo entre 5 y 24.

Conclusiones

Se estudiaron las estructuras de distintas redes de colaboración científica a través de la relación de coautoría en los artículos argentinos publicados en revistas locales e internacionales. El índice de productividad respondió a la ley de Lotka con un valor similar al observado en las publicaciones biomédicas. El tamaño de la red de colaboración fue inferior al esperado, debido posiblemente a la reducida ventana de tiempo del estudio. La distancia promedio entre los autores fue mayor que la esperada, lo que indicaría una estructura inadecuada de conexiones y colaboración entre los investigadores.

REV ARGENT CARDIOL 2009;77:487-492.

Palabras clave > Bibliometría - Cardiología - Argentina

INTRODUCCIÓN

Una red social es un conjunto de personas y relaciones que interactúan en un medio determinado y que suele representarse en un grafo como una colección de puntos y vectores interconectados. El estudio de las redes sociales tiene una base teórica sólida y un interés práctico importante, dada la posibilidad de

determinar los distintos patrones de vinculación e influencia entre las personas, las vías de dispersión de una enfermedad o las formas de difusión de la información en una comunidad. (1-6) En este contexto, la estructura de las redes de colaboración científica se ha estudiado recientemente en distintas disciplinas. (7, 8) Por red de colaboración científica se entiende la vinculación de coautoría que existe entre los diferen-

MTSAC Miembro Titular de la Sociedad Argentina de Cardiología

† Para optar a Miembro Titular de la Sociedad Argentina de Cardiología

¹ Sociedad Argentina de Cardiología, Argentina

² Bioestadística, Facultad de Ciencias Biomédicas, Universidad Austral, Argentina

tes autores de trabajos de investigación en una disciplina, así como la cantidad de colaboradores por autor y el número de artículos publicados. De esta forma, dos autores se consideraran conectados si han colaborado en un mismo artículo. La construcción de gráficos de colaboración entre investigadores se ha llevado a cabo para varias disciplinas a través del análisis de grandes bases de datos como *Medline* y mediante el auxilio de *software* especializado capaz de interconectar los distintos autores y coautores en una gran red de cooperación científica. Las características de estas estructuras y los parámetros derivados de estas redes se utilizan para describir diferentes índices de productividad y colaboración, como son el promedio de publicaciones y colaboradores por autor, el promedio de autores por artículo, el exponente tau (τ) de productividad, el tamaño de la red, las distancias media y máxima entre autores y el llamado "componente gigante" de la red. (7)

El uso de estas redes no es nuevo en el campo de la bibliometría, ya que también se han empleado para vincular las citas bibliográficas entre los distintos artículos publicados. (9-11) Sin embargo, últimamente se ha prestado más atención a las redes de coautoría dentro de una comunidad académica, ya que éstas podrían ser más representativas de la estructura del conocimiento de dicha comunidad. (8)

El objetivo de este trabajo fue describir la estructura de una red de colaboración científica en la Argentina a partir del análisis de coautoría de los artículos cardiológicos publicados durante 2007 en *Medline* y en dos revistas locales de la especialidad.

MATERIAL Y MÉTODOS

Se realizó una búsqueda bibliográfica de todos los artículos cardiológicos publicados por autores argentinos durante 2007. La búsqueda incluyó los artículos aparecidos en *Medline* durante ese período y la totalidad de los publicados en dos revistas de cardiología locales: *Rev Argent Cardiol* y *Rev Fed Arg Cardiol*. Para la identificación en *Medline* de los artículos provenientes de la Argentina se emplearon los criterios de restricción << país >> como filiación del primer autor o como lugar de publicación, << cardio* >> con la misma función de restricción, para seleccionar los centros de los cuales provenían los artículos, y la función de límite temporal para abarcar sólo el año 2007, teniendo en cuenta las diferencias de idioma (Argentina AND Argentine). Una estrategia complementaria de búsqueda para delimitar las patologías cardiovasculares incluyó los términos *Mesh* incorporados en el Apéndice. Por su parte, la pesquisa en las dos revistas locales se realizó en forma manual. Se consideró todo tipo de artículo, ya sea original de investigación clínica o básica aplicada, caso clínico, editorial y revisión, todos del ámbito de la cardiología.

La recolección de datos incluyó el nombre de cada autor y coautores por artículo a fin de determinar la cantidad de colaboradores por autor y de autores y coautores por trabajo. El número de artículos por autor se representó en un gráfico de distribución de frecuencias en escala logarítmica y el índice de productividad se expresó a través del exponente tau (τ) de la ley potencial de Lotka, (12) de acuerdo

con la siguiente relación: frecuencia = n^τ artículos^{- τ} . A partir de esta información se crearon distintas redes de colaboración en las que cada autor o coautor de un artículo se relacionó con otros colaboradores con los que compartían otros artículos publicados, hasta crear un entramado de nodos y relaciones. Se construyeron sendas redes para los autores que publicaron en *Rev Argent Cardiol*, *Rev Fed Arg Cardiol* y *Medline*, por separado y en conjunto. Las redes se graficaron mediante el uso del algoritmo Kamada-Kawai (13) en el *software* Pajek. (14) Este algoritmo ubica los nodos y las conexiones de forma tal que se solapen lo menos posible en el gráfico bidimensional, a fin de producir imágenes interpretables. Tras la construcción de las redes, se calcularon los siguientes parámetros:

- **Componente gigante de la red.** En las redes con un número pequeño de conexiones, los individuos suelen pertenecer a pequeñas islas de colaboración o comunicación. A medida que la cantidad de conexiones aumenta, se genera un gran conjunto de individuos conectados que se denomina componente gigante de la red. Cuanto mayor es el componente gigante, mejor es la interconexión entre los autores.
- **Diámetro del componente gigante.** La máxima separación entre dos autores, también llamada diámetro del componente gigante, mide la distancia más grande que se debe recorrer para conectar dos individuos, en términos de cantidad de nodos interpuestos. Obviamente, este cálculo excluye a los individuos no conectados.
- **Distancia o grados promedio de separación.** A partir del cálculo de la distancia mínima entre todos los pares de autores de la red, se obtiene la distancia promedio de separación para el conjunto de los individuos conectados.

RESULTADOS

Durante 2007 se publicaron 239 artículos cardiológicos provenientes de la Argentina, distribuidos según se indica en la Tabla 1. El promedio de artículos por autor fue similar en las dos revistas locales y algo inferior para las revistas incluidas en *Medline*, posiblemente debido a la mayor exigencia de la publicación en las revistas indexadas. Con respecto a la productividad científica de los autores argentinos, el exponente τ mostró en todos los casos un valor $> 2,0$, lo que indica que la producción no está dominada por pocos individuos con un gran número de colaboradores, sino por múltiples autores con pocos colaboradores. De todas maneras, se observó una tendencia en la que la *Rev Argent Cardiol* tuvo la menor diversidad de autores (mayor dominancia de pocos) y las publicaciones en *Medline* la mayor variedad (menos dominancia de pocos). Lógicamente, la tendencia de crecimiento del exponente τ es inversamente proporcional al promedio de artículos por autor. En la Figura 1 se muestra el número de artículos publicados por autor durante 2007, junto con las curvas de productividad esperadas a partir del exponente τ ; en todos los casos se observó una buena correlación entre la producción observada y esperada, basada en los resultados de los coeficientes de determinación R^2 . El promedio de autores/coautores por artículo, así como el promedio de colaboradores por autor fue más alto en los artículos

Tabla 1. Resumen de los resultados de las redes de colaboración científica de los artículos cardiológicos publicados en *Rev Argent Cardiol* y *Rev Fed Arg Cardiol* y *Medline* durante 2007

	RAC	FAC	Medline
Total de artículos	100	41	98
Total de autores	80	37	92
Promedio de artículos por autor	1,24	1,20	1,12
Exponente tau (τ)	2,78	3,08	3,45
Total de coautores (colaboradores)	287	70	477
Promedio de autores por artículo	3,67	3,60	6,51
Promedio de colaboradores por autor	3,59	2,60	4,88
Parámetros de la red			
Tamaño del componente gigante	131	14	505
Porcentaje del componente gigante	35,7%	13,1%	58,2%
Distancia promedio	4,0	1,5	3,3
Distancia máxima	11	5	10

RAC: *Rev Argent Cardiol*. FAC: *Rev Fed Arg Cardiol*.

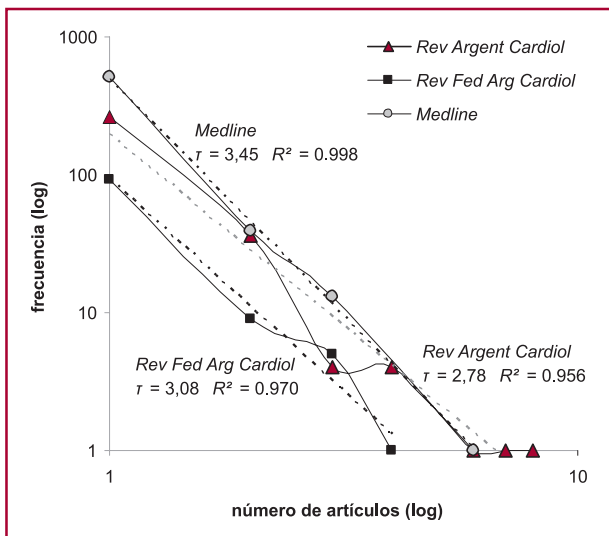


Fig. 1. Número de artículos publicados por investigador durante 2007 en *Medline*, *Rev Argent Cardiol* y *Rev Fed Arg Cardiol*, junto con las curvas de productividad científica esperadas obtenidas a partir del exponente tau (τ), en líneas punteadas paralelas a cada curva observada. Los coeficientes R expresan las correlaciones entre las curvas de productividad observadas y esperadas.

publicados en *Medline* que en las dos revistas locales. Esto podría deberse a que la mayoría de los artículos argentinos en *Medline* son originales y de investigación; en cambio, las revistas locales incluyen más editoriales y revisiones, habitualmente firmadas por uno o dos autores.

Al construir las redes individuales y calcular los parámetros en la Tabla 1, se observó una variabilidad en el tamaño del componente gigante, que en el caso de *Medline* fue el más grande. Por su parte, las dis-

tancias promedio o grados de separación entre dos autores y los diámetros de las redes fueron similares para *Rev Argent Cardiol* y *Medline*, lo que indicaría un grado de conexión similar en ambas redes. En el caso de *Rev Fed Arg Cardiol*, el reducido tamaño del componente gigante se debería al menor número de artículos registrados en el año, lo que disminuye la probabilidad de conexión entre los autores.

La topología de la red en conjunto que incluyó todos los artículos cardiológicos argentinos publicados durante 2007 se muestra en la Figura 2. La red construida en Pajek con el algoritmo Kamada-Kawai representa el componente gigante y algunos componentes accesorios separados del anterior en la región NE y SE de la figura (conglomerados de nodos desvinculados del componente gigante). Los distintos colores señalan los autores que publicaron exclusivamente en *Medline*, en *Rev Argent Cardiol* o en *Rev Fed Arg Cardiol*, y, por otro lado, aquellos que publicaron en dos o más de ellos. El componente gigante de esta red creció hasta incluir el 65,8% de los nodos en una estructura semejante a dos hemisferios unidos centralmente por autores con publicaciones predominantemente en *Rev Argent Cardiol* y/o *Medline*.

DISCUSIÓN

La relación de coautoría en los artículos publicados en revistas científicas permite analizar en alguna medida la estructura de colaboración dentro de una comunidad académica. En este trabajo se describieron las características de una red de colaboración científica en cardiología a partir de los artículos publicados tanto en revistas locales no indexadas como en las pertenecientes a *Medline* durante un año.

Cuando se analizó la distribución del número de artículos publicados por autor, se encontró que ésta se adaptaba bien a la conocida ley de Lotka de pro-

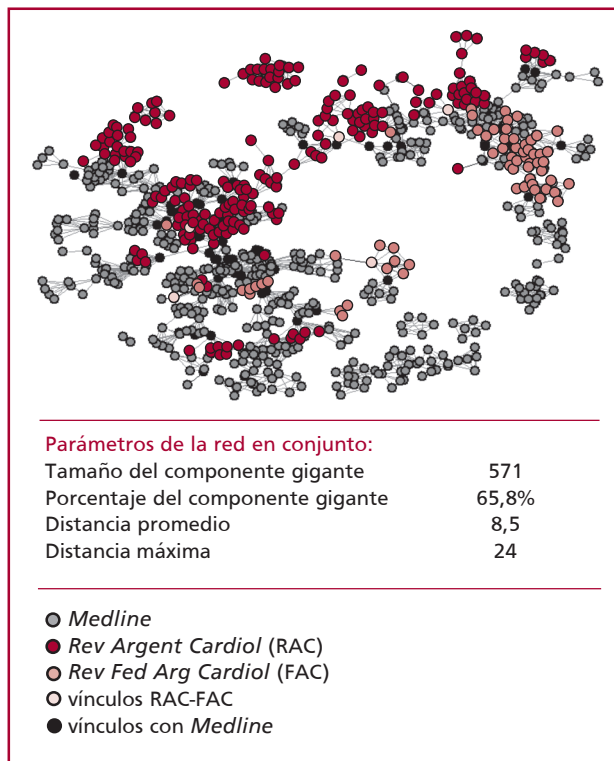


Fig. 2. Estructura de la red de colaboración científica que incluye los autores que publicaron artículos cardiológicos en *Medline*, *Rev Argent Cardiol* y/o *Rev Fed Arg Cardiol* durante 2007. La red que representa al componente gigante y algunos componentes secundarios se construyó con el algoritmo Kamada-Kawai (13) en el software Pajek. (14) Cada color corresponde a los autores que publicaron exclusivamente en *Medline*, en *Rev Argent Cardiol* o en *Rev Fed Arg Cardiol* y, por otro lado, aquellos que publicaron en dos o más de ellos (vínculos).

ductividad científica. Comparativamente, al estudiar las redes de colaboración científica en distintas disciplinas, Newman (7) halló que el exponente τ en las llamadas “ciencias duras” se encontraba cercano a 1; en cambio, en el área de la biomedicina era 2,5, similar al que generan las conexiones de la *World Wide Web*. El valor de $\tau = 2$ señala un límite que divide los comportamientos diferentes de las redes. Por un lado, con una $\tau < 2$, la principal característica de la red es la de estar dominada por un pequeño número de autores con una gran cantidad de colaboradores; mientras que una red con $\tau > 2$ está dominada por muchos autores con pocos colaboradores. Por ejemplo, la red de colaboración de todo *Medline* tiene un $\tau = 2,5$, la de SPIRES de física un $\tau = 1,03$ y la base NCSTRL de ciencias de la computación un $\tau = 1,3$. (7)

El tamaño de la red de colaboración y en especial el del componente gigante son muy importantes para evaluar el nivel de conexión entre los investigadores. Es deseable que el componente gigante englobe a la mayoría de los autores, ya que podría ser un buen índice de colaboración científica. Cuando se compa-

ran los tamaños de los componentes gigantes de nuestro estudio con los referidos en la bibliografía por otros autores, en este último caso se observa que el porcentaje habitual ronda entre el 82% y el 92%, (8) muy por encima de los de la Tabla 1 (13,1% a 58,2%). Esto se debe a que existe una relación directa entre el tamaño del componente gigante y el número total de artículos considerados en el análisis, o sea, la ventana de tiempo elegida para el cálculo. En este sentido, en la Figura 2 se observa el crecimiento del componente gigante hasta el 65,8% después de sumar las tres redes. De la misma manera, si se considerara una ventana de tiempo más larga, el número de artículos aumentaría, y la probabilidad de que se generen más conexiones entre los autores también, con el consiguiente crecimiento del componente gigante. En la Figura 2 se observan además otros componentes secundarios separados del componente gigante, que contienen un escaso número de autores y conexiones.

Los estudios recientes sobre redes describieron la importancia de la distancia promedio entre los nodos. (8) La red de *Medline* estudiada por Newman (7) en el período 1995-1999, con más de 1.500.000 autores, tenía una distancia promedio de 4,6. Comparativamente, la red de la Figura 2 tiene una distancia media de casi el doble (8,5), lo que hablaría en contra de una estructura adecuada de conexiones y colaboración entre los investigadores. Es importante destacar que tanto la red de Newman como la de la Figura 2 tienen el mismo diámetro o distancia máxima (24). La conformación espacial del componente gigante de la Figura 2 podría explicar la excesiva distancia promedio hallada, ya que la red parece estar conformada por dos hemisferios unidos por un istmo central de un solo camino, por el que forzosamente tienen que pasar todas las conexiones de un extremo al otro de la red.

Sin duda, el concepto de colaboración científica se extiende más allá de la publicación en conjunto, ya que existen relaciones informales de comunicación entre los investigadores que no necesariamente culminan en la coautoría de un artículo, lo que suele denominarse “colegas invisibles”. De todas formas, las redes de coautoría descritas en este trabajo son un buen punto de partida para conocer cuantitativamente la estructura de colaboración científica en una comunidad.

Algunas de las limitaciones de este estudio se relacionan con la ventana de tiempo elegida y con el hecho de haber excluido del análisis algunos índices más complejos usados para evaluar redes sociales, como son los coeficientes de conglomerado (*clustering*), centralidad (*betweenness*) y variación (*assortativity*). Otra futura dirección de la investigación no llevada a cabo en este trabajo es el estudio de los cambios de la estructura de la red en el tiempo; en particular, el crecimiento preferencial con incorporación de nuevos coautores a los de los nodos más conectados. (9, 15)

CONCLUSIONES

Se estudió la estructura de distintas redes de colaboración científica en la cardiología argentina, a través de la relación de coautoría en los artículos publicados en revistas locales e internacionales. El índice de productividad por autor respondió a la ley de Lotka y tuvo un valor similar al observado corrientemente en las publicaciones biomédicas. El tamaño de la red de colaboración medido a través del componente gigante fue inferior al esperado, hecho que podría corresponder a la reducida ventana de tiempo usada en el estudio. Por último, la distancia promedio entre los autores fue mayor que la esperada, lo que indicaría una estructura inadecuada de conexiones y colaboración entre los investigadores.

SUMMARY

Scientific Collaboration Networks in Argentine Cardiology Publications

Background

The structure of scientific collaboration networks has been recently studied in different disciplines. The analysis of large data bases using specialized software capable of connecting the different authors and coauthors in a large network of scientific collaboration has enabled the construction of collaboration graphs between investigators in several disciplines. The use of these networks is not new in the field of bibliometrics; however, attention has been recently focused on academic co-authorship networks which might be more representative of the structure of knowledge of an academic community.

Objective

To describe the structure of scientific collaboration networks in Argentina based on co-authorship network analysis of the articles published in the field of cardiology during 2007.

Material and Methods

We conducted a bibliographic search of Argentine papers published in the field of cardiology. Data was retrieved from *Medline* and from two local journals: *Rev Argent Cardiol* and *Rev Fed Arg Cardiol*. Collaboration networks between authors were constructed using the Kamada-Kawai algorithm included in the Pajek software.

Results

Mean papers per author ranged from 1.12 to 1.24, the exponent *tau* of productivity was 2.78 to 3.45, mean authors per paper from 3.60 to 6.51, and collaborators per author ranged between 2.60 and 4.88. The construction of collaboration networks showed that the size of the giant component was between 13.1% and 65.8%, the mean distance between authors was 1.5 to 8.5 and the maximum distance was 5-24.

Conclusions

The structures of different scientific collaboration networks based on co-authorship in Argentine papers published in local and international journals were studied. The productivity index followed Lotka's law with a value that was similar to the one reported in biomedical publications. The size of the collaboration network was smaller than expected,

probably due to the short period of the study. The mean distance between authors was greater than we expected, indicating an inadequate structure of connections and collaboration between investigators.

Key words > Bibliometrics - Cardiology - Argentina

BIBLIOGRAFÍA

1. Girvan M, Newman MEJ. Community structure in social and biological networks. *Proc Natl Acad Sci USA* 2002;99:8271-6.
2. Newman MEJ. The structure and function of complex networks. *SIAM Review* 2003;45:167-256.
3. Nowak MA, Sigmund K. Evolution of indirect reciprocity. *Nature* 2005;437:1291-8.
4. Bearman PS, Moody J, Stovel K. Chains of affection: the structure of adolescent romantic and sexual networks. *Am J Sociol* 2004;110:44-91.
5. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *N Engl J Med* 2007;357:370-9.
6. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. *N Engl J Med* 2008;358:2249-58.
7. Newman MEJ. The structure of scientific collaboration networks. *Proc Natl Acad Sci USA* 2001;98:404-9.
8. Newman MEJ. Coauthorship networks and patterns of scientific collaboration. *Proc Natl Acad Sci USA* 2004;101:5200-5.
9. Grossman JW. The evolution of the mathematical research collaboration graph. *Congressus Numerantium* 2002;158:202-12.
10. Barabási AL, Jeong H, Ravasz E, Neda Z, Schuberts A, Vicsek T. Evolution of the social network of scientific collaborations. *Physica A* 2002;311:590-614.
11. Price DJ. Networks of scientific papers. *Science* 1965;149:510-5.
12. Lotka AJ. The frequency distribution of scientific productivity. *J Wash Acad Sci* 1926;16:317-23.
13. Kamada T, Kawai S. An algorithm for drawing general undirected graphs. *Information Processing Letters* 1989;31:7-15.
14. Batagelj V, Mrvar A. Pajek – analysis and visualization of large networks. En: Junger M, Mutzel P, editors. *Graph drawing software*. Berlin: Springer; 2003. p. 77-103.
15. Newman ME. Clustering and preferential attachment in growing networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 2001;64(2 Pt 2):025102.

APÉNDICE

Términos Mesh incorporados en la búsqueda para delimitar los artículos relacionados con patología cardiovascular

("Cardiomyopathies"[Mesh] OR "Cardiomyoplasty"[Mesh] OR "Cardiology"[Mesh] OR "Cardiomyopathy, Dilated"[Mesh] OR "Diagnostic Techniques, Cardiovascular"[Mesh] OR "Cardiology Service, Hospital"[Mesh] OR "Cardiomyopathy, Hypertrophic"[Mesh] OR "Pregnancy Complications, Cardiovascular"[Mesh] OR "Cardiomyopathy, Hypertrophic, Familial"[Mesh] OR "Chagas Cardiomyopathy"[Mesh] OR "Cardiopulmonary Resuscitation"[Mesh] OR "Cardiovascular Surgical Procedures"[Mesh]) OR ("Cardiotonic Agents"[Mesh] OR "Cardiovascular Infections"[Mesh] OR "Cardiovascular Abnormalities"[Mesh] OR "Tuberculosis, Cardiovascular"[Mesh] OR "Cardiovascular Diseases"[Mesh] OR "Cardiovascular Agents"[Mesh] OR "Cardiopulmonary Bypass"[Mesh] OR "Cardioplegic Solutions"[Mesh] OR

“Cardiomyopathy, Restrictive”[Mesh] OR “Syphilis, Cardiovascular”[Mesh] OR “Cardiomyopathy, Alcoholic”[Mesh] OR “Shock, Cardiogenic”[Mesh] OR “Heart Arrest, Induced”[Mesh] OR (“Electrocardiography”[Mesh] OR “Defibrillators, Implantable”[Mesh] OR “Atrial Natriuretic Factor”[Mesh] OR “Electric Countershock”[Mesh]) OR (“Myocarditis”[Mesh] OR (“Hypertension”[Mesh] OR “Hypertension, Pregnancy-Induced”[Mesh] OR “Hypertension, Renovascular”[Mesh] OR “Hypertension, Renal”[Mesh] OR “Hypertension, Pulmonary”[Mesh]) OR (“Arrhythmia”[Mesh] OR “Anti-Arrhythmia Agents”[Mesh] OR “Arrhythmia, Sinus”[Mesh]) OR (“Pacemaker, Artificial”[Mesh]) OR

(“Myocardial Ischemia”[Mesh] OR “Electrocardiography, Ambulatory”[Mesh] OR “Myocardial Reperfusion Injury”[Mesh] OR “Coronary Disease”[Mesh] OR “Angina Pectoris”[Mesh]) OR (“Heart Rupture, Post-Infarction”[Mesh] OR “Myocardial Infarction”[Mesh] OR “Pulmonary Embolism”[Mesh]) OR (“Electrophysiology”[Mesh] OR (“Angioplasty, Balloon, Laser-Assisted”[Mesh] OR “Angioplasty, Transluminal, Percutaneous Coronary”[Mesh] OR “Angioplasty, Balloon”[Mesh] OR “Angioplasty, Laser”[Mesh]) OR (“Heart Valve Diseases”[Mesh] OR “Pulmonary Valve Insufficiency”[Mesh]) AND (“Country” [Affiliation]) AND (“Date”[EDAT])