

Betegségek kapcsolatrendszerének hálózatai

Machalik Károly¹, Dr. Fogarassyné dr. Vathy Ágnes²

¹ Pannon Egyetem, Rendszer- és Számítástudományi Tanszék,
machalik@dcs.uni-pannon.hu

8200 Veszprém Egyetem utca 10.

² Pannon Egyetem, Rendszer- és Számítástudományi Tanszék,
vathy@dcs.uni-pannon.hu

8200 Veszprém Egyetem utca 10.

Összefoglaló: A betegségek kapcsolatrendszerének populációsintű elemzése rejtett összefüggéseket tárhat fel a betegségek összefüggésére, kialakulására vonatkozóan. A szövédménykutatás egyik eszköze az a hálózatelméleti alapokon nyugvó megközelítés, amely gráfok segítségével szemlélteti és írja le a betegségek kapcsolatrendszerét. Jelen publikációban azt a módszertant és kapcsolódó szoftveralkalmazást mutatjuk be, amely az egészségügyi finanszírozási adatbázis adatai alapján képes betegséghálózatokat generálni és összehasonlítani.

Bevezetés

Magyarországon az egészségügyi intézményeknek az ellátási kötelezettségük mellett jelentési kötelezettségük is van, mivel az állami költségvetési szervek utófinanszírozás keretében térítik meg az ellátó intézmények költségeit. Ezen adatszolgáltatási kötelezettség révén jön létre az egészségügyi finanszírozási adatbázis, amely tartalmazza a páciensek személyes adatait, az alkalmazott egészségügyi ellátás kódját és időpontját, valamint az ellátáshoz kapcsolódó betegségek BNO azonosítóját. Az egészségügyi finanszírozási adatbázis ily módon egy országos méretű elektronikus egészségügyi lánymonitornak tekinthető, amely kellően reprezentálja a magyar lakosság egészségügyi adatait és számos kutatás alapját adja [1,2]. Mindemellett, az adatbázisban tárolt megbetegedések megfelelő alapot biztosítanak ahhoz, hogy elkészítsük a magyar lakosság betegséghálózatát és ezen hálózatot tudományos kutatások során elemezzük.

Módszer

A finanszírozási adatbázisban tárolt megbetegedési adatok egy-egy páciensre vonatkozóan nyújtanak információt a páciens betegségeire vonatkozóan. A rendelkezésre álló adatok azonban nem elemezhetők megfelelő adatelőkészítési műveletek elvégzése nélkül. Az adatelőkészítés során foglalkozni kell az adatok tisztításával, konvertálásával, illetve szűrésével is, hogy csupán a releváns, valós megbetegedési adatokat használjuk fel elemzéseink során. Ezen műveletek közül külön kiemelendő az adatok szűrésének fontossága, mely során gondoskodnunk kell például arról, hogy a beküldő iránydiagnózisok ne kerüljenek a betegséghálózatok készítésekor felhasználásra.

A megfelelő adatelőkészítést követően rendelkezésre álló adatok a páciensek és a betegségek kapcsolatát írják le, amely kapcsolatrendszer egy olyan páros gráfként képzelhetjük el, ahol a páciensek alkotják az egyik, a betegségek pedig a másik objektumhalmazt. Ezen páros gráfból mátrixműveletek segítségével két unáris projekció hozható létre [3]. Egyrészt elkészíthetjük a páciensek kapcsolatrendszerét leíró beteghálózatot, másrészt létrehozhatjuk a betegségek közti összefüggéseket feltáró betegséghálózatot.

A páciens-betegség páros gráfból képezhető betegségekre vetített projekció tehát egymódusú betegséghálózatot eredményez. A hálózatában az egyes csúcspontok a betegségeket reprezentálják és két betegség között akkor jön létre él, ha legalább egy olyan páciens tartalmaz a kiinduló adathalmaz, aki mindkét betegségben szenved. Amennyiben különböző betegcsoportokra készítjük el ily módon a betegségek hálózatát, akkor a betegcsoportokra jellemző speciális hálózatok összehasonlíthatóvá válnak, s ezáltal az egyes betegcsoportokra speciálisan jellemző tipikus betegségek is feltárhatók.

Az így kialakítható betegséghálózatokban az egyes betegségek közötti összefüggések jellemzésére 3 numerikus mérőszámot definiáltunk, melyek révén a betegségek

összefüggésének relevanciája több aspektus szerint is értékelhetővé válik. Ezek a mérőszámok a következők:

- **abszolút gyakoriság:** Az élekre jellemző abszolút gyakoriság megadja, hogy az él által összekötött csomópontok által reprezentált betegségek hány páciens esetében fordulnak elő együtt a vizsgált populációban.
- **relatív gyakoriság:** A relatív gyakoriság az abszolút gyakoriság értékének és a vizsgált populációban/betegcsoportban megjelenő összes kapcsolat számának aránya.
- **feltételes gyakoriság:** A feltételes gyakoriság az abszolút gyakoriság értékét a szóban forgó betegségek közül legalább az egyik betegségben szenvedő páciensek számához arányosítja.

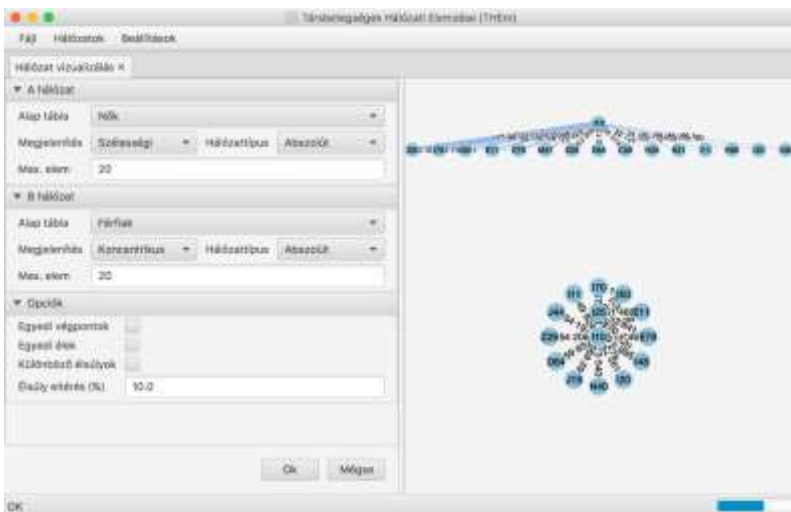
Az adatelemzések során a vizsgálati szempontok alapján számos különféle betegséghálózatot hozhatunk létre. Ezen betegséghálózatok összehasonlítása érdekes összefüggésekre/eltérésekre hívhatja fel a kutatók figyelmét. Ily módon összehasonlíthatóvá válnak például a különféle betegcsoportok betegséghálózatai. Ezen betegséghálózatok összehasonlítása lehetővé teszi például a betegségek megjelenésének betegcsoportokhoz való kapcsolását. További érdekes adatelemzések alapját adhatja a hálózatok időbeni változásának vizsgálata is.

A hálózatok összehasonlítása történhet vizuálisan és numerikusan is. Az elkészült hálózatok megfelelő vizualizálásával kiemelhetők a releváns összefüggések vagy különbözőségek. A vizuális kiemelés mellett fontos szempont azonban, hogy a hálózatok hasonlóságát konkrét mérőszám segítségével is ki lehessen fejezni. Kutatásaink során erre célra a Spearman-féle rangkorrelációs, numerikus értéket használjuk, melyet a hálózatokban megjelenő élek rangsorszámain értelmezünk.

Alkalmazási példa

A hálózatok dinamikus generálására és felhasználóbarát elemzéséhez egy kifejezetten erre a célra fejlesztett alkalmazást

hoztunk létre (1. ábra). Az alkalmazás három fő funkciója: az elemzendő betegcsoportok előállítása, a rájuk jellemző hálózatok automatikus generálása megjelenítése, valamint ezen hálózatok összehasonlítása.



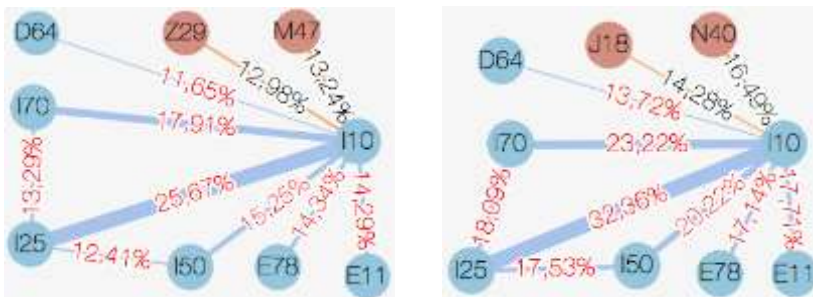
1. ábra A hálózatgeneráló és elemző alkalmazás képernyője

A betegcsoportok előállításakor hétféle adatszűrő áll a felhasználó rendelkezésére, melyek segítségével pontosan definiálható a vizsgálandó betegcsoport. A meghatározott betegcsoport adatai alapján az alkalmazás automatikusan létrehozza a betegcsoport betegséghálózatát. A legenerált hálózatok elmenthetők a további esetleges későbbi felhasználás céljából.

A hálózatok vizualizációja során az élek jellemzésére választhatunk az abszolút, relatív és feltételes gyakoriság mérőszámok közül. Az alkalmazás képes egy, vagy két hálózat együttes megjelenítésére is. Két hálózat megjelenítése esetén a hálózatok közti eltéréseket az alkalmazás grafikusán kiemeli. Az eltérések kiértékelése történhet az új csúcsok, új élek, valamint az élsúlyok alapján. Az élsúlyok eltérését az

alkalmazás a kiválasztott mérőszám alapján értékeli ki és a felhasználó által megadott küszöbértéket meghaladó eltéréseket grafikusan kiemeli.

A 2. ábra az ezen alkalmazás segítségével generált hálózatokra mutat példát. Az ábra bal oldalán egy nagyméretű férfiak alkotta betegcsoport betegség-hálózatának, jobb oldalán pedig ugyanilyen szempontok alapján válogatott női betegcsoport betegség-hálózatának 10 legnagyobb élsúlyal bíró részhálózata tekinthető meg.



2. ábra Példahálózatok. bal oldal: női betegcsoport, jobb oldal: férfi betegcsoport

A hálózatok csúcspontjain a betegségek – három karakterre aggregált – BNO kódja látható, az élek súlyozásaként pedig a betegségkapcsolatok relatív gyakorisága jeleníti meg. A két hálózaton összehasonlításakor megfigyelhetjük, hogy a leggyakoribb betegség-betegség kapcsolat mindkét hálózatban azonos betegségek között jön létre (I25: Idült ischaemiás szívbetegség és I10: Elsődleges magasvérnyomás-betegség), azonban ezen betegségek együttese előfordulásának relatív gyakorisága már jelentős eltérést mutat a két csoport esetében. A női betegcsoport esetében csupán a betegek negyedénél, míg a férfiak esetében a betegek harmadánál található meg egyszerre a magas vérnyomás és az idült ischaemiás szívbetegség. Az ábrán piros színnel vannak kiemelve azon csúcspontok, amelyek nem jelennek meg a másik betegcsoport

10 leggyakoribb élet tartalmazó részhálózatban. Nőknél a spondylosis (M47) és az egyéb profilaktikus eljárások szükségessége (Z29), míg a férfiaknál a prostata túltengés (N40) és a tüdőgyulladás (J18) jelenik meg új csúcsként a hálózatban. A két teljes hálózat hasonlóságát mérő Spearman-féle rangkorrelációs érték (0,7152) közepesen erős hasonlóságot mutat.

Konklúzió

A bemutatott módszertan és alkalmazás hasznos segítséget nyújthat az orvosszakmai kutatások terén. Az alkalmazás segítségével egyrészt igazolhatók az ismert és tapasztalati megfigyelések, másrészt új összefüggések is felfedezhetők. A vizuális kiemelés nagyban segíti az orvosszakmai kutatók munkáját, mivel ennek segítségével a rejtett összefüggések nem csak a numerikus mérőszámok összehasonlításával tárhatók fel, hanem vizuálisan is szemléltethetők.

Köszönetnyilvánítás

A publikációt és a kapcsolódó kutatásokat a VKSZ_12-1-2013-0012 azonosítójú "Világszínvonalú Intelligens és Inkluzív Egészségügyi Információs és Döntéstámogató Keretrendszer (Analytic Healthcare Quality User Information) kutatása" című projekt keretében Magyarország Kormánya támogatta.

- [1] Vassy Zsolt, Kósa István, Vassányi István. „Stabil anginás betegutak klaszterelemzése.” In: XXVIII. Neumann Kollokvium kiadványa. Veszprém, 2015. nov. 20-21, ISBN 978-615-5036-10-1, 7-10 o.
- [2] Király Ferenc, Nemes Attila, Vassányi István, Kósa István. „A szívkatóteres laboratóriumoktól mért földrajzi távolság hatása az iszkémiás szívbetegek gyanújával ellátásra került betegek ellátási útjára.” IME XIII. évfolyam, 3. szám, 2014. április, ISSN 1588-6387, 15-18. o.
- [3] Barabási Albert-László, „A hálózatok tudománya”, Libri Könyvkiadó, 2016