

ROC analízis alkalmazása

Fazekasné Kis Mária

Debreceni Egyetem Agrártudományi Centrum,
Agrárgazdasági és Vidékfejlesztési Intézet,
Agrárinformatikai és Alkalmazott Matematikai Tanszék, Debrecen

ÖSSZEFOGLALÁS

A ROC (Receiver Operating Characteristic) analízis főként a külföldi szakirodalom szerint széles körben alkalmazott módszer a diagnosztikai tesztek hatékonyságának elemzésére. A dolgozat ismerteti a ROC analízis alapelemeit és a ROC görbe készítését, valamint bemutat néhány diagnosztikai vizsgálati eredményt reprezentáló ROC ábrát.

SUMMARY

According to publishings in foreign scientific journals, ROC (Receiver Operating Characteristic) analysis is a widely used method for analysing the diagnostic utility of clinical laboratory tests. In this paper, we explain the basic principles of ROC analysis and produce ROC curves, as well as demonstrate some ROC curves, which represent the results of diagnostic tests.

BEVEZETÉS

A diagnosztikai tesztek hatékonyságának jellemzésére többféle statisztikai módszer alkalmazható. Gyakorlatban fontos a különböző diagnosztikai tesztek hatékonyságának összehasonlítása is. A ROC (Receiver Operating Characteristic) analízis kezdeti felhasználásai az 1950-es években történtek, a radarjel észlelésével kapcsolatban. Az elmúlt évtizedben a módszer alkalmazásáról főként a nemzetközi szakirodalomban egyre gyakrabban olvashattunk (Berbaum et al., 2000; Davelaar et al., 1998; Ignjatovics et al., 2000; Linnet, 1988; Leung et al., 1989; Viitala et al., 1998; Zweig és Campbell, 1993). A módszert sokszor felhasználták a kémiai tesztek hatékonyságának elemzésére, radiológiai képfeldolgozáshoz kapcsolódóan, valamint az orvosi döntéshozatalban is (Erkel és Pattyanama, 1998; Lusted, 1960; Lusted, 1971). Viszonylag ritkán hazánkban is alkalmazták ezt az elemzési módot (Boda, 2000; Fazekasné, 1999; Fazekasné és Góth, 1999).

A ROC elemzéssel egyrészt a diagnosztikai tesztek hatékonyságáról kapunk információt, másrészt jó összehasonlításra ad alkalmat egy ábrán, ha egy betegség diagnosztizálására többféle diagnosztikai teszt eredménye is rendelkezésre áll. A dolgozat áttekintést ad a ROC analízisről, a ROC ábra értelmezéséről, alkalmazásainak előnyeiről.

ROC ANALÍZIS

A tesztek diagnosztikai hatékonyságának méréséhez tradicionálisan a diagnosztikai szenzitivitás (érzékenység) és a diagnosztikai specificitás (specifikusság) fogalmak kapcsolódnak (Linnet, 1988; Zweig és Campbell, 1993). A

diagnosztikai hatékonyság megítéléséhez beteg és egészséges (a vizsgált betegségben nem szenvedők) csoportokban vizsgálják a teszt eredményeit, megállapítják, hogy pozitív vagy negatív. Az 1. táblázat összefoglalja a lehetséges eseteket.

1. táblázat

A diagnosztikai tesztek eredményei

	Teszt(1)	
	Negatív(2)	Pozitív(3)
Egészségesekre(4)	TN(6)	FP(7)
Betegekre(5)	FN(8)	TP(9)

Table 1: The results of diagnostic tests
test(1), negative(2), positive(3), healthy(4), patients(5), true positive(6), false positive(7), false negative(8), true positive(9)

Az angol nyelvű szakirodalomban szokásos rövidítések jelentése:

TP= igaz pozitív (valódi pozitív teszt esetek száma)

TN= igaz negatív (valódi negatív teszt esetek száma)

FP= hamis pozitív (tévesen pozitív teszt esetek száma)

FN= hamis negatív (tévesen negatív teszt esetek száma)

„igaz pozitív” – a teszt által helyesen betegnek minősítettek száma, „igaz negatív” – a teszt által helyesen minősített egészségesek száma, „hamis pozitív” – a teszt által tévesen minősített egészségesek száma, „hamis negatív” – a teszt által tévesen minősített betegek száma.

Ezek alapján:

Diagnosztikai szenzitivitás= $TP/(TP+FN)$

Diagnosztikai specificitás= $TN/(TN+FP)$

Így a diagnosztikai szenzitivitás a betegek számának arányaiból, a diagnosztikai specificitás az egészségesek számának arányaiból határozható meg.

Az ideális jó diagnosztikai teszt eredményeinek eloszlásában nincs átfedés a beteg és az egészségesek csoportjában. A diagnosztikai szenzitivitás és specificitás is 100%-os. A legtöbb diagnosztikai teszt eredményeinek eloszlásában átfedést láthatunk a beteg és az egészségesek eredményei között. Alacsonyabb döntési szintet választva a hamis negatív eredmények száma csökken, de nő a hamis pozitívok száma (nagyobb a diagnosztikai szenzitivitás). Növelve a döntési szint értékét, növekszik a hamis negatív esetek száma, és csökken a hamis pozitívoké (nagyobb a diagnosztikai

specifitása). Összefoglalva: fordított irányú kapcsolat van a diagnosztikai szenzitivitás és specifitása között.

A teszt eredmények szemléltetésére, főleg a nemzetközi szakirodalomban, alkalmazott módszer a ROC analízis. A ROC görbét egységnégyzetben ábrázolják, melynek x-tengelyén az (1-specifitása), az y-tengelyén a diagnosztikai szenzitivitása értékeit jelenítik meg. A négyzet jobboldalán feltüntetethetők az (1-diagnosztikai szenzitivitása) értékeit, a négyzet felső tengelyén pedig a diagnosztikai specifitása értékeit is. A ROC görbe az alkalmazott teszt eredményeinek átfedését ábrázolja a betegek és az egészségesek csoportjainak adataiból, a döntési szintek teljes tartományában. A görbe minden pontja egy diagnosztikai szenzitivitása/diagnosztikai specifitása pontpárt jelenít meg egy döntési szinten. A teszt diagnosztikai hatékonysága akkor a legjobb, ha a diagnosztikai szenzitivitása és diagnosztikai specifitása is maximális, ebben az esetben a ROC görbe áthalad az egységnégyzet bal felső csúcsán. Hamis pozitív és hamis negatív eredmények ilyenkor nem adódnak, a betegek és az egészségesek eredményeit jellemző Gauss-görbék nem fedik egymást. Ha a teszt a vizsgált betegség szempontjából nem tud különbséget tenni, akkor a ROC görbe pontjai az egységnégyzet bal alsó és jobb felső csúcsát összekötő diagonális egyenes körül helyezkednek el.

A ROC görbe alkalmazható több diagnosztikai teszt összehasonlítására is, mikor ugyanazon betegség diagnosztizálására többfajta laboratóriumi teszt eredményei ismertek. Ilyenkor a különböző tesztek eredményei egy ROC ábrán jeleníthetők meg, a görbék egymáshoz és az egységnégyzetbeli elhelyezkedésük alapján hasonlíthatók össze.

A laboratóriumi teszt diagnosztikai hatékonysága kifejezhető a ROC görbe alatti terület mérőszámával. A mérőszám illetve a görbe alatti terület akkor maximális (=1), mikor az alkalmazott teszt a betegek és az egészségesek teszt eredményeit „teljesen jól” megkülönbözteti. A ROC görbe alatti terület számértékét jó közelítéssel megadja a diagnosztikai szenzitivitása/diagnosztikai specifitása pont párok alatti trapézok területeinek összege. A diagnosztikai hatékonyság megítélésénél figyelembe kell venni a ROC görbék alakját is, mivel lehet két görbe alatti terület azonos, de a görbék „alakja” lényegesen eltérő.

DIAGNOSZTIKAI TESZTEK EREDMÉNYEINEK ÖSSZEHAJONLÍTÁSA

A diagnosztikai tesztek eredményeinek szemléltetése történhet gyakorisági hisztogramokon. Az alábbi ábrákon gyakorisági hisztogramokon láthatóak a vér szérumban kataláz enzim eredmények 229 akut szívinfarktusból beteg (1. ábra) és 111 egészséges (2. ábra) egyén csoportjaiban. A két csoport eredményeit jellemző Gauss-görbe átfedést mutat. A betegek csoportjában 80-90 U/ml értéknél van csúcs, az egészségeseknél a leggyakoribb érték 50-60 U/ml.

1. ábra: A vérszérumban kataláz enzim értékek a betegek csoportjában

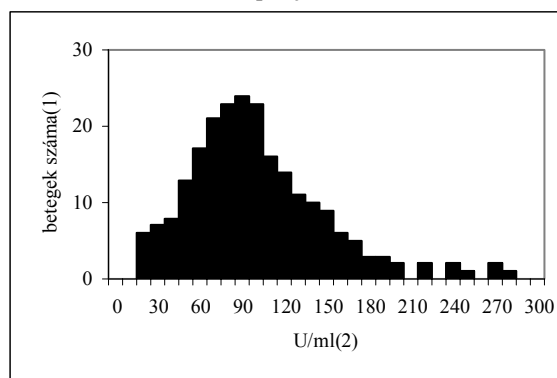


Figure 1: The value of catalase enzymes for the group of patient number of patients(1), U/ml(2)

2. ábra: A vérszérumban kataláz enzim értékek az egészségesek csoportjában

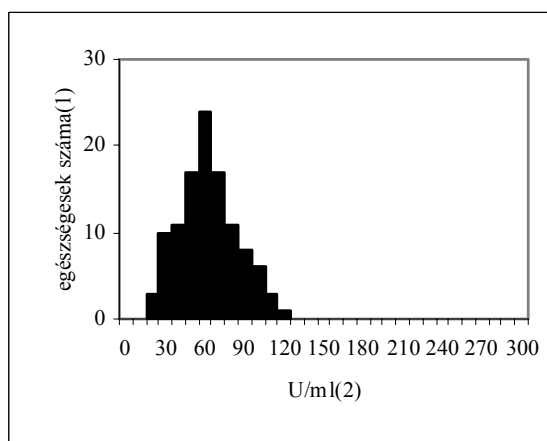


Figure 2: The value of catalase enzymes for the group of healthy number of healthy(1), U/ml(2)

Az egészségesek vér szérumban kataláz aktivitása a (20, 120) U/ml tartományba, a betegeké pedig a (20, 280) U/ml tartományba estek.

A gyakorisági hisztogramok jól mutatják az egyes mérési intervallumokban a betegek és az egészséges egyének teszt eredményeinek számát. Viszont a diagnosztikai teszt hatékonyságáról semmilyen információt nem nyújtanak.

A következő ábrák (3-4. ábra) 105 akut epehólyag gyulladásban beteg és 114 egészséges egyén csoportjaiban szintén a vér szérumban kataláz enzim aktivitását mutatják. Mindkét ábrán az 50-60 U/ml értéknél van a legnagyobb gyakoriság. Az egészségesek eredményei a (0, 130) U/ml tartományba, a betegeké a (10, 330) U/ml tartományba estek.

A ROC analízis alkalmazására először olyan példát mutatunk be, mikor ugyanazt a diagnosztikai tesztet (vér szérumban kataláz aktivitás mérést) alkalmaztuk két különböző betegség diagnosztizálására.

3. ábra: A vérérum kataláz enzim aktivitás értékek akut epehólyag gyulladásban betegek csoportjában

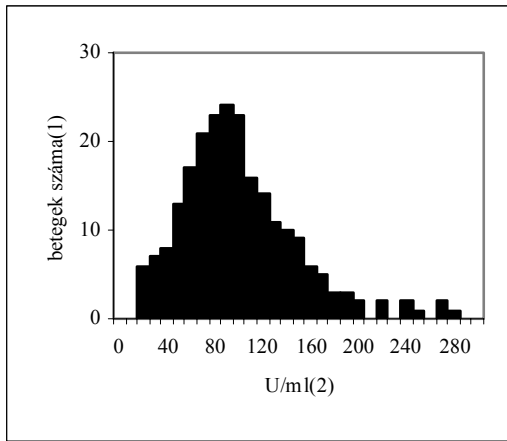


Figure 3: The activity of blood serum catalase for acute cholecystitis patients number of patients(1), U/ml(2)

4. ábra: A vérérum kataláz enzim aktivitás értékek akut epehólyag gyulladásban egészségesek csoportjában

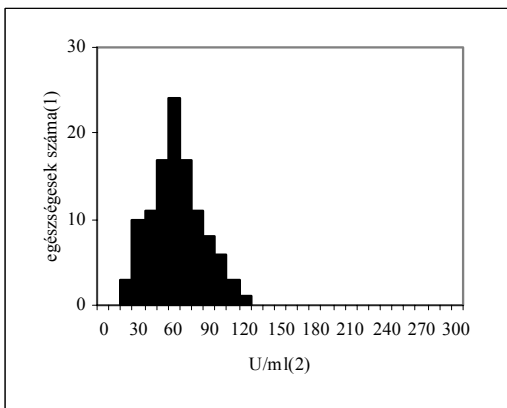


Figure 4: The activity of blood serum catalase for acute cholecystitis patients who are otherwise healthy number of healthy(1), U/ml(2)

Az előzőekben említett betegségek és számadatok alapján a ROC görbék pontjait az ismertett módon meghatároztuk és elkészítettük a ROC ábrákat. Az 5. ábra mutatja az akut szívinfarktusban szenvedők és egészségesek csoportjainak ROC ábráját. A görbe alatti terület mérőszáma megadja diagnosztikai teszt hatékonyságát, erre 0,843 érték adódott.

5. ábra: ROC analízis akut szívinfarktus vizsgálatára

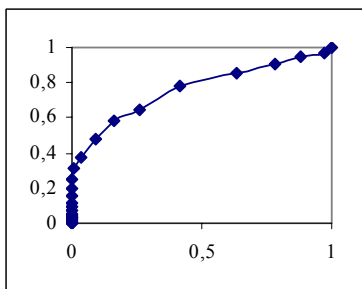


Figure 5: ROC analysis for acute myocardial infarction

A 6. ábrán az akut epehólyag gyulladásban betegek és ehhez képest egészségesek adataiból készült ROC görbe látható. A görbe alatti területre 0,652 értéket kaptunk, a görbe inkább az egységnégyzet diagonálisához van közelebb, mint a négyzet bal felső csúcsához. A görbe alakja alapján azt mondhatjuk, hogy alacsonyabb diagnosztikai szintnél érzékenyebb a meghatározás, mint nagyobb értékeknél. A két ROC görbe területi mérőszáma és a görbék egységnégyzetbeli elhelyezkedésük összehasonlításából arra következtethetünk, hogy a vér szérum kataláz enzim aktivitás az akut szívinfarktus diagnosztizálására nagyobb hatékonyságú, mint az akut epehólyag gyulladásra. Az utóbbi betegség gyanúját célszerű más módon is megvizsgálni.

6. ábra: ROC analízis akut epehólyag gyulladás vizsgálatára

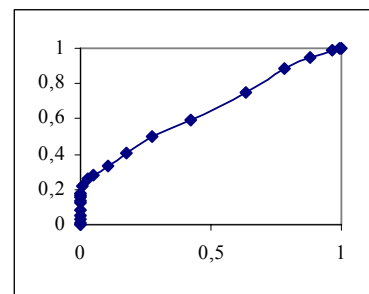


Figure 6: ROC analysis for acute cholecystitis

Klinikai laboratóriumokban a ROC analízis másik jól alkalmazható módja, mikor több különböző diagnosztikai teszt hatékonyságát hasonlítjuk össze egy betegség diagnosztizálásához. A 7. ábrán a ROC görbék az akut hasnyálmirigy gyulladás diagnosztizálásához alkalmazott lipáz és α -amiláz tesztek eredményeit hasonlítják össze. A ROC görbék alapján is látható, hogy a lipáz enzim jobb hatékonyságú. Ezt igazolják a görbe alatti területek mérőszámai is, mivel az α -amiláz görbe alatti területe 0,5295, ugyanez a lipáz esetében 0,9882.

7. ábra: ROC analízis akut hasnyálmirigy gyulladás esetén α -amiláz (●) és lipáz (◆) enzimek vizsgálatával

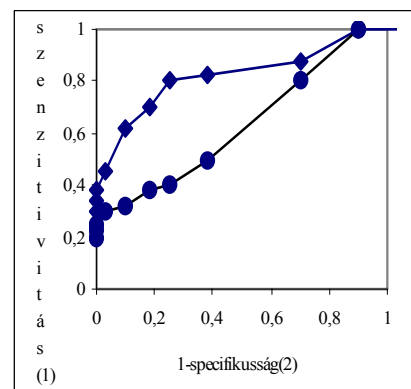


Figure 7: ROC analysis for acute pancreatitis with α -amilase (●) and lipase (◆) enzyme tests sensitivity(1), 1-specificity(2)

KÖVETKEZTETÉSEK

A ROC analízis alkalmas a kiválasztott betegség diagnosztizálásához alkalmazott többféle teszt összehasonlítására, azt vizsgálva, melyiknek jobb a diagnosztikai hatékonysága. Ez a módszer a laboratóriumokban másrészt segítheti egy diagnosztikai módszer hatékonyságának megítélését többféle betegség megállapításában. A ROC görbe az elemzési eredményeket szemléletesen mutatja, a diagnosztikai teszt hatékonyságának számszerű értékét a görbe alatti terület adja, ezek alapján jól alkalmazható a döntés hozatalban.

A ROC ábra egyszerű, grafikus módon becslést ad a diagnosztikai teszt hatékonyságára. Az elemzéshez rendelkezésre álló adatok gyakoriságától független, nem szükséges, hogy a beteg és az

egészséges csoport egyedeinek száma azonos legyen. Nem kell megadni egy speciális döntési szintet, mivel a ROC ábra magában foglalja a döntési szintek teljes spektrumát. Több teszt szemléletes összehasonlítására is alkalmazható egy ábra, míg ha a hagyományos hisztogramokkal vagy vonaldiagramokkal történik a teszt eredmények prezentálása, akkor több ábrát is kell készíteni, ha a mértékegységek eltérőek. A ROC görbe készítésekor az adatok csoportosítása elmarad, szemben a hisztogramok alkalmazásával. A ROC ábra előállításakor kapunk információt a diagnosztikai teszt szenzitivitásáról (érzékenységről) és specifitásáról (specifikusságról), ami nem mondható el a teszt eredmények hagyományos szemléltetési módjai esetében.

IRODALOM

- Berbaum K. S.-Dorfmann D. D.-Franken E. A.-Caldwell Rt. (2000): Proper ROC analysis and joint ROC analysis of the satisfaction of search effect in chest radiology. *Acad Radiol*, 7. 11. 945-958.
- Boda K.-Kozinszky Z.-Balogh Zs. (2000): Többváltozós regressziós modellek alkalmazása orvosi kutatási adatokra statisztikai programrendszerekkel. XXII. Neumann Kollokvium. A számítástechnika orvosi és biológiai alkalmazásai. 61-64.
- Davelaar E. M.-Kamp G. J.-Verstraeten R. A. et al. (1998): Comparison of seven immunoassays for the quantification of CA 125 antigen in serum *Clin Chem.*, 44. 1417-1422.
- Erkel A. R.-Pattynama P. M. (1998): Receiver operating characteristic (ROC) analysis: Basic principles and applications in radiology. *European Journal of Radiology*, 27. 88-94.
- Fazekasné Kis M. (1999): A ROC analízis oktatása az orvosdiagnosztikai laboratóriumi analitikus szakon. *Informatika a Felsőoktatásban '99*. 228-233.
- Fazekasné Kis M.-Góth L. (1999): A ROC analízis alkalmazása az enzim diagnosztikában. *Klin. Kísér. Lab. Med.*, 26. 184-186.
- Ignjatovic S.-Majkic-Singh N.-Mitrovic M.-Gvozdenovic M. (2000): Biochemical evaluation of patients with acute pancreatitis. *Clin Chem., Lab Med.*, 38. 11. 1141-1144.
- Kazmierczak S. C.-Lente F.-Hodges E. D. (1991): Diagnostic and prognostic utility of phospholipase. A activity in patients with acute pancreatitis: comparison with amylase and lipase. *Clin Chem.*, 37. 3. 356-360.
- Keim V.-Teich N.-Reich A.-Fiedler F.-Mössener J. (1997): Polyclonal pancreatic elastase assay is superior to monoclonal assay for diagnosis of acute pancreatitis. *Clin Chem.*, 43. 12. 2339-2344.
- Leung F. Y.-Galbraith L. V.-Jablonsky G. et al. (1989): Re-evaluation of the diagnostic utility of serum total creatine kinase and chreatine kinase-2 in myocardial infarction. *Clin Chem.*, 35. 1435-1440.
- Linnet K. (1988): A review on the methodology for assessing diagnostic tests. *Clin. Chem.*, 34. 7. 1379-1386.
- Lusted L. B. (1960): Logical analysis in roentgen diagnosis. *Radiology*, 74. 178-193.
- Lusted L. B. (1971): Signal detectability and medical decision making. *Science*, 171. 1217-1219.
- Viitala K.-Landesmaki K.-Niemela O. (1998): Comparison of the axis %CDT TIA and CDtect method as laboratory tests of alcohol abuse. *Clin Chem.*, 44. 1209-1215.
- Zweig M.-Campbell G. (1993): Receiver-Operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin. Chem.*, 39. 4. 561-577.