

HODNOTENIE ÚSPEŠNOSTI LABORATÓRNYCH METÓD

**BRANKO BALLA^a, JÁN MOCÁK^a,
ELEONORA VARMUSOVÁ^b, DENISA KAVKOVÁ^b
a IVAN TUDÍK^b**

^a*Katedra analytickej chémie, Fakulta chemickej a potravinárskej technológie, Slovenská technická univerzita, 812 37 Bratislava, ^bÚstav tuberkulózy a respiračných chorôb, Kvetnica, 058 87 Poprad
mocak@cvt.stuba.sk*

Došlo 25.10.02, prepracované 15.2.03, prijaté 18.3.03.

Kľúčové slová: miera citlivosti, miera špecifickosti, účinnosť, hranica rozhodovania, ROC-krivky

Úvod

Chemometrické spracovanie výsledkov laboratórnych metód vytvára nové možnosti ako podporiť proces diagnostikovania v laboratórnej medicíne a v klinickej chémii. Táto oblasť laboratórnych meraní priniesla v ostatnom čase nové prvky do teórie analytickej chémie, vďaka čomu to, čo je obsahom tohto prehľadného referátu, platí nielen o výsledkoch analytických a biochemických stanovení v klinickej praxi, ale je použiteľné všeobecnejšie pre hodnotenie akýchkoľvek analytických meraní¹ či laboratórnych meraní vôbec. Táto práca si kladie za úlohu vysvetliť význam a objasniť použitie najdôležitejších pojmov používaných na charakterizovanie laboratórnych výsledkov, ako sú napr. miera citlivosti a miera špecifickosti², kritické hodnoty (hranice rozhodovania), alebo ROC-krivky, ako aj ukázať možnosti, ktoré poskytujú dostupné softvéry a odkázať na najvhodnejšiu literatúru pre podrobnejšie štúdium.

Binárne rozhodovanie a kontingenčná tabuľka

Častou úlohou pri posudzovaní laboratórnych meraní je vyhodnotenie ich úspešnosti alebo výkonnosti a zistenie, do akej miery možno laboratórne výsledky použiť na spoľahlivú výpoveď o študovaných objektoch. To je obzvlášť významné pri použití laboratórnych meraní na klinicko-diagnostické účely, kde napr. rozpoznanie stavu ochorenia na základe úrovne hodnôt tej-ktorej posudzovanej laboratórnej metódy (v klinickej praxi sa viac používa termín laboratórny parameter) môže priniesť včasné zistenie zdravotného rizika alebo aspoň môže ušetriť príslušnú osobu od zložitejších a invazívnejších diagnostických vyšetrení.

Laboratórna metóda pri binárnom rozhodovaní poskytuje binárnu výpoveď, existujúcu obvykle v tvare áno/nie, prítomný/neprítomný alebo pozitívny/negatívny. Vo vše-

becnosti sa uvažujú prelínajúce sa populácie dvoch kategórií, ktorým síce v zásade zodpovedajú odlišné hodnoty laboratórnych výsledkov (malé vs. veľké, alebo naopak), ale bez úplnej, ideálnej diferenciácie, takže spraviť deliacu čiaru, kde končí jedna a začína druhá populácia, nie je jednoduchá úloha. Celý problém je potrebné chápať z pravdepodobnostného hľadiska. Náhodnosť je príznačnou črtou nielen samotného analytického či biochemického merania, ale napr. v prípade laboratórnych meraní pre klinicko-diagnostické účely je ich náhodná zložka značne ovplyvnená aj zložitou procesov v ľudskom organizme, ktorý v tomto prípade predstavuje extrémne zložitú maticu. Naviac ho v tomto prípade značne ovplyvňuje preanalytická fáza laboratórnej metódy (angl. tiež: laboratory test, assay), kde zarúčenie jednotných podmienok odberu vzorky, vrátane správania sa vyšetřovaného subjektu napr. 24 hodín pred odberom, je stav, ku ktorému sa realita len viac-menej približuje.

V tejto práci sa pre názornosť a jednoduchosť budú uvažovať nasledujúce dve kategórie: *a*) chorí (ill) s veľkosťou súboru n_1 a *b*) zdraví (healthy) s veľkosťou súboru n_2 . Rozdelenie subjektov do týchto dvoch skupín sa má vykonať vždy nezávislým spôsobom, napr. iným, vysoko hodnoverným laboratórnym meraním. Napr. na dôkaz malígneho ochorenia sa používa jednoznačný histologický nález. Zatriedenie vyšetřovanej osoby do skupiny chorých alebo zdravých však v tomto prípade možno s vysokou úspešnosťou predpokladať na základe hodnôt analyzovaných tumorových markerov³⁻⁶, čo je omnoho jednoduchšie a pre vyšetřovaného priateľnejšie. Na výpočet jednotlivých charakteristík skúmaných laboratórnych metód sa potom používajú rôzne druhy porovnania medzi výstupom – výsledkom metódy a správnym zatriedením skúmaného objektu (v tomto prípade subjektu, vyšetřovanej osoby) do jednej či druhej skupiny podľa nezávislého jednoznačného testu. Testovací súbor dát použitých na hodnotenie úspešnosti laboratórnych metód obsahuje len dáta, ktorých príslušnosť do jednej z dvoch uvažovaných kategórií (skupín, tried) je známa. Ide o tzv. „learning set“ (učebný súbor, t.j. dáta, podľa ktorých sa vybraná chemometrická technika učí ako diskriminovať predmetné kategórie) a z hľadiska chemometrickej metodiky ide o tzv. „supervised learning“ (riadené učenie), pri ktorom zatriedenie objektov či subjektov do kategórií je vopred známe. Naproti tomu zadelenie objektov či subjektov do tried pri metodikách typu „unsupervised learning“ (neriadené učenie) nie je apriórne známe (napr. pri analýze hlavných komponentov).

Základom výpočtov pri hodnotení úspešnosti laboratórnych metód je kontingenčná tabuľka (tab. I). Pri konštrukcii kontingenčnej tabuľky sa pre meranie danou laboratórnou metódou najprv určí, resp. odhadne kritická hodnota (nazývaná aj hranica rozhodovania, angl. decision level alebo cut-off value), podľa ktorej sa vykoná zatriedenie do príslušnej skupiny, *i* alebo *h*. Potom sa zatriedenie získané laboratórnou metódou porovná so správnym, nezávisle získa-

ným zatriedením, z čoho rezultuje pre každé meranie jedna zo štyroch možností: *a*) true positive, *tp*, ak výstup laboratórnej metódy sa zhoduje so správnym zatriedením do skupiny s pozitívnym nálezom, *b*) false positive, *fp*, ak výstup laboratórnej metódy poskytuje zatriedenie do skupiny s pozitívnym nálezom, ktoré nie je správne, *c*) true negative, *tn*, ak výstup laboratórnej metódy sa zhoduje so správnym zatriedením do skupiny s negatívnym nálezom, a *d*) false negative, *fn*, ak výstup laboratórnej metódy poskytuje zatriedenie do skupiny s negatívnym nálezom, ktoré nie je správne.

Tabuľka I

Kontingenčná tabuľka 2×2 – posúdenie výsledku laboratórnej metódy

Výsledok laboratórnej metódy	Správny výsledok (t.j. správna diagnóza)		Celkom
	ill (chorí)	healthy (zdraví)	
pozitívny	<i>tp</i>	<i>fp</i>	<i>tp</i> + <i>fp</i>
negatívny	<i>fn</i>	<i>tn</i>	<i>fn</i> + <i>tn</i>
Celkom	<i>tp</i> + <i>fn</i> = n_1	<i>fp</i> + <i>tn</i> = n_2	n_1 + n_2

Spôsoby hodnotenia úspešnosti laboratórnych metód

Charakterizovanie výsledkov zhromaždených v kontingenčnej tabuľke umožňujú nasledujúce, vzájomne sa dopĺňajúce charakteristiky^{1,7-9}: citlivosť *Se* (senzitivita – z angl. sensitivity), špecifickosť *Sp* (specificity) a účinnosť *E* (efficiency), ktoré sú definované v rovniciach (1)–(3):

$$Se = tp / (tp + fn) = tp / n_1 = 1 - \beta = P[p|i] \quad (1)$$

$$Sp = tn / (tn + fp) = tn / n_2 = 1 - \alpha = P[n|h] \quad (2)$$

$$E = (tp + tn) / (tp + fn + fp + tn) = (tp + tn) / (n_1 + n_2) \quad (3)$$

Symbody n_1 a n_2 označujú počet subjektov v prvej, resp. druhej kategórii, α a β označujú štatistickú chybu prvého druhu (pravdepodobnosť, že neplatí nulová hypotéza, ak je v skutočnosti správna), resp. chybu druhého druhu (pravdepodobnosť, že sa prijme nulová hypotéza, ak je v skutočnosti nesprávna). Chyby prvého a druhého druhu sa vyskytujú nielen pri rozmanitom štatistickom testovaní, ale majú napr. základný význam pri určovaní násobiaceho koeficientu v prípade medze detekcie alebo medze stanovenia¹⁰⁻¹².

Podmienené pravdepodobnosti, ktoré sa nachádzajú v rov. (1), (2), ako aj ďalej v rov. (6) a (7), majú nasledujúci zmysel: $P[p|i]$ je pravdepodobnosť, že daný subjekt bude laboratórnou metódou posudzovaný ako pozitívny (*p*), ak patrí naozaj do skupiny chorých (*i*); $P[n|h]$ je pravdepodobnosť, že daný subjekt bude laboratórnou metódou posudzovaný ako negatívny (*n*), ak patrí naozaj do skupiny zdravých (*h*); $P[i|p]$ je pravdepodobnosť, že daný subjekt bude diagnostikovaný ako chorý (*i*), ak jeho laboratórny test (výsledok laboratórnej metódy) je

pozitívny (*p*); $P[h|n]$ je pravdepodobnosť, že daný subjekt bude diagnostikovaný ako zdravý (*h*), ak jeho laboratórny test je negatívny (*n*).

Na tomto mieste treba upozorniť, že v analytickej chémii dávnejšie zavedená citlivosť má aj iný význam, keďže v oblasti kvantitatívnej analýzy je definovaná ako smernica kalibračnej závislosti, vypočítaná pomocou derivácie $\partial S / \partial c$, kde *S* je signál a *c* označuje koncentráciu¹³⁻¹⁵. Preto AOAC označuje veličinu vyjadrenú v rov. (1) ako miera citlivosti (angl. sensitivity rate)². Aj špecifickosť sa niekedy používa v kvantitatívnej analýze v inom význame, a to na charakte-

rizovanie vplyvu interferencií v multikomponentnom analyte¹⁵, preto AOAC veličinu vyjadrenú v rov. (2) označuje ako miera špecifickosti (angl. specificity rate)². Tam, kde k nejednoznačnosti nemôže dôjsť, používajú sa však stručnejšie termíny citlivosť a špecifickosť, čo je bežné napr. v klinickej chémii.

V odbornej literatúre sa vyskytujú aj ďalšie charakteristiky výsledku laboratórnej metódy¹: pozitívny pravdepodobnostný pomer (angl. positive likelihood ratio, *PLR*) negatívny pravdepodobnostný pomer (angl. negative likelihood ratio, *NLR*), definované v rov. (4) a (5), pozitívna prediktívna hodnota (angl. positive predictive value, *PPV*) a negatívna prediktívna hodnota (angl. negative predictive value, *NPV*), definované v rov. (6) a (7), kde sa rôznym spôsobom kombinujú pomery, v ktorých sa vyskytujú *tp*, *fp*, *tn* a *fn*:

$$PLR = (tp/n_1) / (fp/n_2) = Se / (1 - Sp) \quad (4)$$

$$NLR = (tn/n_2) / (fn/n_1) = Sp / (1 - Se) \quad (5)$$

$$PPV = tp / (tp + fp) = P[i|p] \quad (6)$$

$$NPV = tn / (tn + fn) = P[h|n] \quad (7)$$

PPV má podobný priebeh ako *Sp*, t.j. hodnoty *PPV* sú v rozsahu $<0;1>$ a s hodnotou príslušnej premennej stúpajú. *NPV* má podobný priebeh ako *Se*, teda hodnoty *NPV* sú v rozsahu $<0;1>$ a s hodnotou príslušnej premennej klesajú. *PLR* má so vzrastom príslušnej premennej stúpajúci trend a jej priebeh je málo hladký (značne zubatý); *NLR* má naopak klesajúci trend. Rozsah hodnôt *PPV*, ako aj *NPV* nie je zhora pevne ohraničený (zdola je nula), takže ich stupnice v porovnaní s citlivosťou a špecifickosťou majú výrazne väčší rozsah.

Hranica rozhodovania a ROC-krivky

Hranica rozhodovania δ , prípadne nazývaná aj ako kritická hodnota (angl. decision limit, cut-off value), definuje hodnotu, nad ktorou sa výstup laboratórnej metódy hodnotí ako pozitívny; hodnoty pod medzou rozhodovania sa hodnotia ako negatívne^{1,9,16}. To však platí len v prípade, ak hodnoty testu vzťahované na populáciu **h** (zdraví) sú menšie oproti hodnotám vzťahovaným na populáciu **i** (chorí), inak sa vykoná presné opačné hodnotenie⁹. Ak by populácie **h** a **i** boli úplne separované, potom by platilo, že $tn = n_2$ ($fp = 0$) a $tp = n_1$ ($fn = 0$). V reálnom živote sú však rozdelenia pravdepodobností obidvoch kategórií vzájomne prekryté a zvolená hodnota δ reguluje hodnoty charakteristík úspešnosti laboratórneho testu, t.j. citlivosti, špecifickosti, účinnosti, prípadne hodnôt *PLR*, *NLR*, *PPV* a *NPV*. Vzrast citlivosti je vždy sprevádzaný poklesom špecifickosti, a naopak. Podobné antagonizmy sú aj v prípadoch *PLR* a *NLR*, ako aj *PPV* a *NPV*.

Správna voľba hranice rozhodovania δ závisí od cieľa laboratórneho merania, t.j. či klasifikácie *fp* sú viac škodlivé než klasifikácie *fn*, alebo naopak. Riziko, resp. cena nesprávnej klasifikácie, by mali byť známe predtým, než sa zvolí spôsob, ako určiť optimálnu hodnotu δ . Najobvyklejšia voľba medze rozhodovania δ je daná hodnotou laboratórneho testu, zodpovedajúcou maximu účinnosti (angl. efficiency).

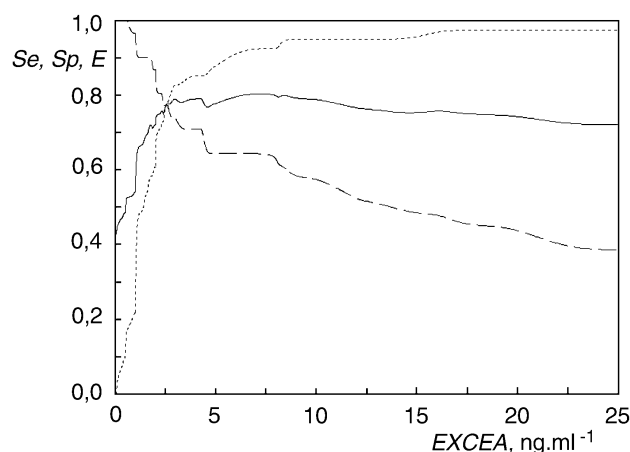
Prediktívnu schopnosť laboratórnej metódy (testu) možno efektívne zviditeľniť pomocou ROC-krivky, ktorá je vynesená v súradniciach *Se* vs. $(1 - Sp)$. Skratka ROC je odvodená od slov receiver operating characteristic, keďže sa pôvodne využívala ako operačná charakteristika rádiológátora¹⁷; až neskôr sa objavili aplikácie ROC v medicíne¹⁸⁻²⁰ a v analytickom kontexte sa ROC vyskytuje len v ostatnej dobe¹. Diagonálna línia na ROC-diagrame predstavuje prípad celkom neúspešnej laboratórnej metódy s nulovou prediktívnou hodnotou, keďže pravdepodobnosť úspešného a neúspešného zatriedenia je zhodne 50 %, čo zodpovedá celkom náhodnej klasifikácii. Vzdialenosť ROC-krivky od diagonálnej línie je preto mierou výkonnosti laboratórnej metódy, čím je vzdialenosť väčšia, tým metóda lepšie odlišuje dve uvažované populácie (zmena od 50 % smerom k 100 %)²¹. V ROC-diagramoch sa ROC-krivky majú v zásade nachádzať v hornom trojuholníku nad diagonálou; nespĺnenie tejto podmienky indikuje nesprávne zadanie dát; presnejšie takýto prípad indikuje plocha pod ROC-krivkou menšia ako 0,5.

Zo štatistického hľadiska ROC-krivky definujú vzťah medzi chybou $(1 - \beta)$ a chybou α , pričom chyba α (chyba I. druhu) reprezentuje frakciu fp/n_1 výsledkov, chyba β (chyba II. druhu) reprezentuje frakciu fn/n_2 výsledkov, takže chyba $(1 - \beta)$ prislúcha frakcii výsledkov tp/n_1 . ROC-krivka ukazuje pre každú možnú hodnotu hranice rozhodovania δ , aké je percento správne diagnostikovaných abnormálnych subjektov (t.j. *tp*) voči percentu nesprávne diagnostikovaných normálnych subjektov (t.j. *fp*)²¹.

Výhodou ROC-analýzy je, že pomocou ROC-kriviek možno znázorniť a porovnať niekoľko laboratórnych metód

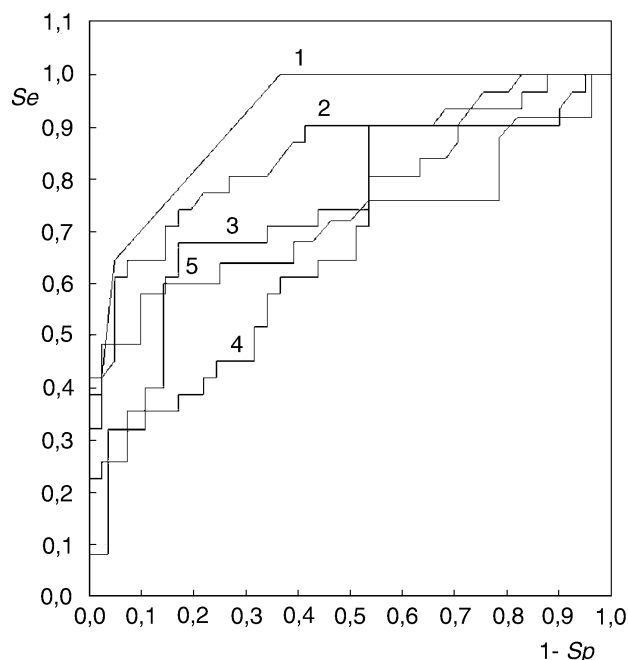
a napokon z nich možno vybrať najlepšiu metódu⁹. Metódy, ktoré lepšie diskriminujú dve predmetné kategórie objektov, resp. subjektov, majú ROC-krivky čo najviac hore a vľavo. Ideálna diagnostická metóda, ktorá by umožnila dokonale diskriminovať normálne a abnormálne subjekty, mala by vykazovať ROC-krivku strmo stúpajúcu hneď od počiatku smerom k ľavému rohu diagramu a potom smerujúcu takmer horizontálne k pravému hornému rohu diagramu²¹. Táto zásada umožňuje vizuálne porovnanie ROC-kriviek, teda kvalitatívny spôsob hodnotenia laboratórnych metód, umožňujúci zostaviť poradie ich výkonnosti. V prípade prekrytých ROC-kriviek je však takáto vizuálna inšpekcia dosť subjektívna.

Hlavným kritériom výkonnosti laboratórnej metódy je plocha *A* vypočítaná pod ROC-krivkou; čím je väčšia, tým je metóda účinnejšia^{9,22-25}. Ak sa do jedného grafu vynesie niekoľko ROC-kriviek, možno jednoduchým porovnaním plochy pod krivkami zistiť, ktoré laboratórne meranie poskytuje najväčšiu diskrimináciu medzi uvažovanými kategóriami, teda v nami uvažovanom prípade medzi skupinami chorých (**i**) a zdravých (**h**). V našich prácach³⁻⁶ sme napr. ukázali, že medzi skúšanými údajmi laboratórnych metód, predstavujúcich tumorové markery na diagnostikovanie malígnych pľúcnych ochorení, najlepšiu informáciu poskytuje premenná CEA vo výpotku (*EXCEA*), potom CYFRA 21-1 vo výpotku (*EXCYF*), ďalej CEA v sére (*SCEA*) a najmenší informačný obsah má CYFRA 21-1 v sére (*SCYF*). Pri



Obr. 1. Grafický priebeh kriviek citlivosti *Se* (---), špecifickosti *Sp* (---) a účinnosti *E* (—) pre premennú *EXCEA*, ktorú predstavuje pľúcny tumorový marker CEA (karcinoembryonálny antigén) v pleurálnom výpotku (*EX*). Vyšetřilo sa 72 osôb, z toho 41 histologicky bez malignity a 31 s histologicky dokázanou malignitou

tomto štúdiu sa použili dva tumorové markery, a to CEA (karcinoembryonálny antigén) a CYFRA 21-1 (cytokeratín fragment), v dvoch pacientom odobratých telesných tekutinách – vo výpotku, čo je kvapalina v oblasti pohrudnice, vyskytujúca sa pri benígnych i malígnych pľúcnych ochoreniach, a v krvi. V názve premennej sme vzorky z výpotku



Obr. 2. ROC-krivky pre 4 premenné predstavujúce pľúcne tumorové markery CEA a CYFRA v pleurálnom výpotku (indikuje počiatková skratka EX) a v sére (počiatková skratka S), ako aj pre PCI multikomponent, získaný lineárnou kombináciou EXCEA, EXCYF, SCEA a SCYF technikou PCA. Vyšetřilo sa 72 osôb, z toho 41 histologicky bez malignity a 31 s histologicky dokázanou malignitou. Plochy pod krivkami A: (1) PCI: $A = 0,9182$; (2) EXCEA: $A = 0,8482$; (3) EXCYF: $A = 0,7549$; (4) SCEA: $A = 0,6869$; (5) SCYF: $A = 0,6864$. Poradie premenných pri hodnote súradnice $(1 - Sp) = 0,3$ je: 1, 2, 3, 5, 4 a pri hodnote $(1 - Sp) = 0,7$ je: 1, 2, 4, 3, 5.

označili ako EX a vzorky v krvnom sére ako S. Na obr. 1 je ukážka priebehu kriviek Se , Sp a E pre EXCEA, na obr. 2 štyri spodné ROC-krivky zodpovedajú uvedeným štyrom premenným. Takto sa dalo zistiť, že medzi individuálnymi údajmi markerov najrelevantnejšiu indikáciu malígneho ochorenia v tomto štúdiu poskytuje CEA vo výpotku (EXCEA). Ešte väčšiu plochu A však poskytuje multikomponent PCI, čo je prvý hlavný komponent zistený technikou PCA (analýzou hlavných komponentov), ktorý bol lineárnou kombináciou troch pôvodných premenných, EXCEA, EXCYF a SCEA, teda obsahuje kolektívnu informáciu najlepších tumorových markerov. Plocha A pod najvyššou krivkou zodpovedajúcou multikomponentu PCI je výrazne väčšia v porovnaní s najlepšou individuálnou premennou EXCEA. Zistili sme tiež, že väčšiu plochu A v porovnaní s premennou EXCEA dáva aj multikomponent DFI, čo je prvá diskriminačná funkcia, vypočítaná ako lineárna kombinácia pôvodných premenných technikou LDA (lineárnou diskriminačnou analýzou), v ktorej sa však lineárna kombinácia premenných vykonáva iným spôsobom než pri PCA.

O charakteristikách úspešnosti laboratórných metód, voľbe hranice rozhodovania, ROC-krivkách a ďalšej problematike

diskutovanej v tejto práci existuje dnes už pomerne veľa literárnych prameňov, ktoré umožnia hlbšie štúdium. Na prvom mieste tu treba spomenúť prehľadné referáty²⁷⁻³¹. Asi najobsažnejším rýchlym literárnym zdrojom je harvardská web stránka K. H. Zou³².

Chemometricky podporovaná predpoveď diagnózy

Pre zvolenú laboratórnu metódu, ktorú možno všeobecne označiť symbolom x , možno výber kritickej hodnoty δ pre premennú x urobiť na základe priebehu kriviek $Se - x$, $Sp - x$ a najmä $E - x$, v tomto často používanom prípade podľa súradnice x maxima účinnosti E . Určenie kritickú hodnotu δ potom možno využiť na chemometrickú predikciu diagnózy u ďalších osôb vyšetřovaných pomocou metódy x . Určenie hodnotu δ možno použiť buď priamo na diagnostikovanie, alebo na porovnanie s uznanou kritickou hodnotou používanou v klinickom laboratóriu, prípadne na jej korekciu. Je však zrejmé, že vykonanie viacerých laboratórných metód s cieľom spresniť diagnózu obvyklým spôsobom, poskytuje len sekvenciu postupne získaných jednotlivých diagnostických informácií. Významný zisk diagnostickej informácie predstavuje určenie kritickej hodnoty z multikomponentu, nájdeného optimálnou lineárnou kombináciou jednotlivých premenných. Treba zdôrazniť, že predikcia diagnózy pomocou multikomponentu je jednoduchá a ľahko vykonateľná v praxi. Poslúži na to nasledujúca príkladová štúdia, v ktorej sa predpoveď pľúcnej malignity vykonáva pomocou prvého hlavného komponentu PCI.

Príkladová štúdia

Technikou PCA (analýza hlavných komponentov) sme zistili rovnicu na výpočet PCI pomocou troch vyššie opísaných zvolených metód, a to EXCEA, EXCYF a SCEA. Premennú SCYF sme vynechali, pretože je najmenej informatívna (na základe ROC-kriviek, obr. 2). Pre testovací súbor 72 pacientov sa v PCA použité tri premenné lineárne kombinovali podľa rovnice získanej vo výstupe použitého softvéru (STATGRAPHICS):

$$PCI = 0,6667 EXCEA + 0,3060 EXCYF + 0,6797 SCEA \quad (8)$$

Z maxima krivky účinnosti pre PCI, t.j. E vs. PCI , sme vo výstupe použitého softvéru (GraphROC for Windows) zistili kritickú hodnotu $-0,442$ na stupnici premennej PCI. Ide o bezrozmernú hodnotu, čo je zapríčinené štandardizáciou, opísanou ďalej. Na predpoveď diagnózy pacienta s ešte neurčenou diagnózou je potrebné vedieť uvedenú kritickú hodnotu a koeficienty v rov. (8). Naviac sú potrebné priemerné hodnoty EXCEA, EXCYF a SCEA v testovacom súbore 72 pacientov a príslušné smerodajné odchýlky, pretože v metóde PCA sa originálne dáta predspracúvajú štandardizáciou, t.j. od jednotlivých laboratórne zistených hodnôt x_i sa odčíta ich priemer $\bar{x}$ a výsledok sa podelí smerodaj-

nou odchýlkou s_x . Hodnoty priemerov a smerodajných odchýlok sa ľahko získajú zo súboru nameraných údajov x_i v tabuľke MS EXCEL alebo v inom tabuľkovom procesore.

Predpoveď diagnózy pre pacienta A

Po štandardizácii nameraných laboratórnych hodnôt sa získali údaje pre pacienta A:

$EXCEA = -0,388$, $EXCYF = -0,162$, $SCEA = -0,234$.

Výpočtom podľa rov. (8) sa získala príslušná hodnota $PCI = -0,467$.

Keďže $-0,467 < -0,442$, diagnostický záver je: benígny.

Predpoveď diagnózy pre pacienta B

Po štandardizácii nameraných laboratórnych hodnôt sa získali údaje pre pacienta B:

$EXCEA = 2,294$, $EXCYF = 0,033$, $SCEA = -0,273$.

Výpočtom podľa rov. (8) sa získala príslušná hodnota $PCI = 1,354$.

Keďže $1,354 > -0,442$, diagnostický záver je: malígny.

Citlivosť, špecifickosť, účinnosť a ďalšie ukazovatele laboratórnych metód, ako aj ROC-krivky, možno vypočítať pomocou viacerých programových balíkov. Sú to hlavne softvéry: a) GraphROC for Windows 2.0 (cit.^{9,33}), b) Analyse-it 1.61 (cit.²¹), c) CBstat 4.1.0 (cit.³⁴), d) MedCalc 6.14 (cit.³⁵), e) SYSTAT 9 (cit.³⁶), f) NCSS 2000 (cit.³⁷) a g) SAS 8.1 (cit.³⁸). Z nich najbližšie k bezprostrednému uplatneniu pri riešení problémov klinickej diagnózy sú GraphROC, Analyse-it, CB-stat a MedCalc. Naproti tomu SYSTAT 9 rieši problémy tohto druhu veľmi všeobecne ako rozlíšenie samotného šumu od signálu so šumom, čo je potenciálne využiteľné na rozmanité aplikácie v analytickej chémii. Modul SIGNAL v SYSTATE umožňuje analyzovať dáta zatriedované do 2 až 11 kategórií, pričom sa používa buď binárna stupnica (pre 2 kategórie) alebo poradová stupnica. Kombinované použitie viacerých softvérových balíkov umožňuje vybrať numerické a grafické výstupy najvhodnejšie pre študovaný problém.

Záver

Táto práca podáva stručný prehľad o základných ukazovateľoch výkonnosti laboratórnych metód, ako sú najmä citlivosť, špecifickosť a účinnosť. Podáva tiež základnú informáciu o voľbe hranice rozhodovania (kritickej hodnoty) a o ROC-krivkách a ich využití na diagnostické účely. Kriticky zostavený výber významných článkov a knižných prác, zaoberajúcich sa uvedenou problematikou, umožní doplniť požadované detaily podľa potreby. Uvádzajú sa tiež najdôležitejšie softvérové balíky, využiteľné na správnu a podrobnú interpretáciu výsledkov laboratórnych analytických a biochemických metód.

Autori ďakujú Slovenskej grantovej agentúre za podporu grantu VEGA1/9129/02.

LITERATÚRA

1. Massart L. M., Vandeginste B. M. G., Buydens L. M. C., De Jong S., Lewi P. J., Smeyers-Verbeke J.: *Handbook of Chemometrics and Qualimetrics*. Part A, kapitola 13 a 16. Elsevier, Amsterdam 1997.
2. AOAC (Association of Official Analytical Chemists): *Definitions and calculations proposed for method performance parameters, 6-12 March 1995*. The Referee.
3. Varmusová E., Mocák J., Balla B., Kavková D., Tudič I.: *5th Symposium with Int. Participation Labkvalita '01, Poprad, 4-6 June 2001*, str. 243. Slov. lek. spoloč., Bratislava 2001.
4. Balla B., Mocák J., Pivovarníková H., Balla J., Kavková D., Varmusová E.: *Clin. Chem. Lab. Sci.* 39, 285 (2001).
5. Kavková D., Varmusová E., Tudič I., Mocák J., Balla B., Berezová M.: *Eur. Respir. J.* 18, 397 (2001).
6. Balla B., Mocák J., Varmusová E.: *Laborat. Diagnostika* 6, 65 (2001).
7. Gerhardt W., Keller H.: *Scand. J. Clin. Lab. Invest.* 1986, Suppl. 181, 1.
8. Galen R. S., Gambino S. R.: *Beyond Normality: The Predictive Value and Efficiency of Medical Diagnoses*. Wiley, New York 1975.
9. Kairisto V., Poola A.: *Scand. J. Clin. Lab. Invest.* 1995, Suppl. 222, 43.
10. Massart L. M., Vandeginste B. M. G., Deming S. N., Michotte Y., Kaufman L.: *Chemometrics: a Textbook*, kap. 7. Elsevier, Amsterdam 1988.
11. Mocák J., Bond A. M., Mitchell S., Scollary G.: *Pure Appl. Chem.* 69, 297 (1997); IUPAC document 550/35/87.
12. Meloun M., Militký J.: *Kompendium statistického zpracování experimentálních dat*. Academia, Praha 2002.
13. Currie L. A.: *Pure Appl. Chem.* 67, 1699 (1995).
14. ISO Standard 8466-1 (1994): *Water quality – Calibration and evaluation of analytical methods and estimation of performance characteristics, Part 1: Statistical evaluation of the linear calibration function*.
15. Vessman J.: *Pure Appl. Chem.* 73, 1381 (2001), IUPAC Recommendations 2001.
16. Bairagi R., Suchindran C. M.: *Ind. J. Stat., Ser. B.* 51, 263 (1989).
17. Egan J. P.: *Signal Detection Theory and ROC Analysis*. Academic Press, New York 1975.
18. Lusted L. B.: *Science* 171, 1217 (1971).
19. McNeil B. J., Adelstein S. J.: *J. Nucl. Medicine* 17, 439 (1976).
20. Swets J. A.: *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*. Lawrence Erlbaum, Mahwah 1995.
21. Analyse-it Software Ltd.: *Manuál k softvéru Analyse-it 1.61*, Leeds 1997–2000. <http://www.analyse-it.com>, staženo 11.12.2000.
22. Hanley J. A., McNeil B. J.: *Radiology* 439, 29 (1982).
23. Hays R. D.: *Appl. Psych. Meas.* 14, 208 (1990).
24. Hilden. J.: *Med. Decision Making* 11, 95 (1991).

25. Zweig M. H., Campbell G.: Clin. Chem. 39, 561 (1993).
26. Emir B., Wieand S., Jung S. H., Ying Z.: Stat. Med. 19, 511 (2000).
27. Hsiao J. K., Bartko J. J., Potter W. Z.: Arch. Gen. Psychiatry 46, 664 (1989).
28. Begg C. B.: Stat. Medicine 10, 1887 (1991).
29. Campbell G.: Stat. Medicine 13, 499 (1994).
30. Centor R. M.: Med. Decision Making 11, 102 (1991).
31. Shultz E. K.: Clin. Chem. 41, 1248 (1995).
32. Zou K. H.: *Receiver Operating Characteristic (ROC) Literature Research*. <http://www.spl.harvard.edu:8000/pages/ppl/zou/roc.html>, staženo 10.9.2002.
33. Kairisto V., Poola A.: *Manuál k softvéru GraphROC for Windows 2.0.*, Turku – Tallinn 1994-96. <http://www.netti.fi/~maxiw>, staženo 28. 05. 2001.
34. Linnet K.: *Manuál k softvéru CBstat 4.1.0.* Risskov 2001. <http://www.cbstat.com>, staženo 8. 6. 2001.
35. Schoonjans F.: *Manuál k softvéru MedCalc 6.14.* Gent 2002. <http://www.medcalc.be>, staženo 10. 9. 2002.
36. SPSS: *Manuál k softvéru SYSTAT 9.* SPSS Science, Chicago 1998. <http://www.systat.com>, staženo 10. září 2002.
37. SAS: *SAS Software Release 8.2.* SAS Institute, Cary 2000. <http://www.sas.com>, staženo 8. 4. 2002.
38. NCSS: *Manuál k softvéru NCSS 2000.* NCSS Statistical Software, Kaysville 2000. <http://www.ncss.com>, staženo 20.7.2000.

B. Balla¹, J. Mocák¹, E. Varmusová², D. Kavková², and I. Tudík² (¹*Department of Analytical Chemistry, Faculty of Chemical and Food Technology, Slovak University of Technology, Bratislava*, ²*Institute of Tuberculosis and Respiratory Diseases, Poprad - Kvetnica, Slovak Republic*): **Assessment of Outcome of Laboratory Methods**

Chemometrical processing of the results of laboratory methods create new possibilities for supporting decisions about appropriate diagnosis in laboratory medicine and clinical chemistry. Used approach, originally applied in medicine, has been recently transferred also into analytical chemistry. Qualitative characteristics of laboratory methods, like Sensitivity, Specificity or Efficiency are used for the selection of the appropriate decision limit, according to which all examined objects or subjects can be classified into chosen two categories (e.g. healthy and ill subjects). Assessment of performance of the compared laboratory methods is made by means of ROC-curves. Utilization of multivariate statistical methods, combining original variables in an optimal way, increases the diagnostic value of laboratory assays. Case study, based on the multicomponent calculation by Principal Component Analysis, shows how the linear combination of the analysed tumor markers increases their diagnostic value for the differential diagnostics of pleural malignancy.