A pajzsmirigy-szcintigráfiás leletező(“Orvosi íródeák” program) klinikai tesztelési tapasztalatai

Az orvosi diktálóprogram tesztelésére a Szegedi Tudományegyetem Orvostudományi Karának Nukleáris Medicina Intézetének orvosait kértük fel. A tesztelésben 5 olyan röntgendiagnosztikai szakember vett részt, akik rendszeresen végeznek pajzsmirigy-szcintigráfiai leletkiértékelést. A tesztelést végző személyeket arra kértük, hogy fejenként legalább 5-5 leletet olvassanak be. Segítségképpen a tesztelés idejére úgy alakítottuk át a program felületét, hogy a rendelkezésünkre álló lelet-adatbázisból véletlenszerű leleteket kínáljon fel – de természetesen a tesztelő személy más, eltérő leletanyagot is bemondhatott. A tesztszemélyeknek lehetőségük volt a bediktált mondatokhoz akár mondatonként is megjegyzéseket fűzni a diktálás hatékonyságára vonatkozólag, de arra is megkértük őket, hogy a tesztelés végén pár mondatban összegezzék benyomásaikat. A tesztelés során bemondott hanganyagot rögzítettük. Ez egyrészt lehetővé teszi majd a hibák forrásának alapos utólagos elemzését. A fontosabb következménye azonban, hogy a felvételek segítségével objektív módon ki tudtuk értékelni a szószintű találati pontosságot százalékos arányban. Ez egyfelől azért érdekes, mert a felhasználók szubjektív benyomása nem feltétlenül arányos a valódi pontossággal, másfelől pedig a rövid kommentárokban nyilván főleg csak a negatív tapasztalataikat jegyezték fel a tesztszemélyek, így ezek alapján könnyen a valósnál negatívabb képet kaphatunk a teszteredményekről. A rögzített hanganyagokon elért találati pontosságot ezért minden esetben kiszámítottuk és a tesztelők szóbeli értékelése után közöljük (adaptáció előtt és után).

Az adaptációs programmodul tesztelése céljából arra kértük a tesztszemélyeket, hogy az adaptációs blokkban is olvassanak be fejenként minimum 5-5 leletet, majd a teszt üzemmódba visszatérve vizsgálják meg, hogy a fő diktálási modulban javult-e a diktálási hatékonyság. A program tesztelése egy 1800KHz-es Pentium IV-es gépen történt, 256Mb memória mellett, egy HAMA 46040-es mikrofonnal.

Az 1. tesztelő személy tapasztalatai:

Az egyik fontos benyomásom a programmal az volt, hogy a rendszer nagyon érzékeny a háttérzajra, például a telefoncsörgésre, vagy ha valaki belép a szobába. Ilyenkor ezekre a zajokra is kiír valamit, amit aztán már nem lehet javítani. Ez komoly problémát okozhatna olyankor, amikor ugyanabban a szobában több leletező is dolgozik egyszerre. Ettől a problémától eltekintve azonban elég meggyőző találati pontossággal működött a rendszer. A rövidebb mondataimat mindig teljesen pontosan eltalálta. Voltak azonban szavak, amiket nem szeretett. Rájöttem például, hogy a „két” helyett kettőt olvasva sokkal jobban eltalálja a 2-es számot. Ezektől az apróságokon kívül számomra elég meggyőző volt a program. Némi gyakorlás, hozzászokás után valószínűleg élesben is lehetne használni. Az orvosi szakkifejezések írásmódját még pontosítani kellene, de ha jól sejtem, ez részletkérdés. A felhasználói felületet jó lenne hozzáigazítani a leletnyilvántartó

rendszerünkhöz, vagy akár esetleg az egész diktálási mechanizmust beleépíteni mondjuk a PACS rendszerbe.

Találati pontosság adaptácó előtt: 89.57% *Adaptáció után: 91.10%*

A 2. tesztelő személy tapasztalatai:

Azt, hogy "a jugulumot" szinte soha nem ismeri fel, 3-t ír helyette.

Gyakran a már jól beírt szót átírja egy másikra, ráadásul olyanra, amit nem is mondok.

A változás helyett sokszor elváltozást ír, a nincs helyett is-t.

"göb" szót nem ismeri fel, erre mindent ír össze-vissza.

Gyakran a lebenyben helyett csak lebeny-t ír.

Kis fújásra, mikrofonhoz érésre gyakran beírja a "nyak" szót.

A "pajzsmirigy" szót sokszor csak sokkal később írja ki, ekkor általában már megismétlem.

Összességében: Elég hosszasan teszteltem a programot. Bizonyos szavakat nagyon nem szeret valamiért. Más mondatokra viszont egész jól működött.

Találati pontosság adaptácó előtt: 83.57% *Adaptáció után: [nem tesztelte]*

A 3. tesztelő személy tapasztalatai:

A rendszer tesztelésekor a felkínált mondatokat igyekeztem beolvasni. Nekem már alapállapotban is jól működött a rendszer, egészen keveset tévesztett. Viszont sokszor nagyon lassan követi a diktálást, és többször ki is dobott a programból vagy hibajelzést adott, ami után be kellett zárni a programot. Megpróbálkoztam a beszélőadaptációs szöveg beolvasásával is, de utána nem tapasztaltam érzékelhető változást, talán azért, mert már alaphelyzetben is jónak találtam a felismerést. A lassú reakcióidőtől eltekintve abszolút használhatónak tűnt a program. Lehet, hogy gyorsabb gép kellene??

Találati pontosság adaptácó előtt: 94.42% *Adaptáció után: 95.74%*

A 4. tesztelő személy tapasztalatai:

A rövid mondataimat felismeri a program, de a hosszúakban gyakran téveszt. Ilyenkor teljesen „eltéved”, és a mondat hátralévő részét elrontja. Nem világos számomra, hogy ilyenkor mit kellene tenni. Nem volt teljesen nyilvánvaló az sem, hogy a rövidítések esetén (pl. UH) a rövidítést vagy a teljes szót kell-e bemondani.

Az adaptációhoz szükséges beolvasandó szöveg elég hosszú (igazság szerint a kért 5 lelet helyett végül 15-öt olvastam végig, további javulás reményében). Viszont utána egész határozottan javult a rendszer, most már a hosszabb mondatokat is sikerül végig beolvasni.

Összefoglalva a tapasztalatimat, az első benyomásom nem volt túl jó. Azonban némi használat után, és főleg az adaptációs folyamat végigcsinálása után megváltozott a véleményem. Most azt gondolom, hogy van gyakorlati potenciál a programban. Bár szükségesnek látszik az adaptációs folyamat végigvitele, azt gondolom, az eredmény

megéri. Az biztos, hogy szokni kell a programot és a használatát. Nem ártana pontosabb leírás arról, hogy a rövidítéseket, latin szavakat meg ilyesmiket hogyan kell beolvasni.

Találati pontosság adaptácó előtt: 76,37% *Adaptáció után: 93,87%*

Az 5. tesztelő személy tapasztalatai:

A nyak szót rengetegszer kiírja olyankor is, amikor pedig nem is diktáltam.

Nem szereti a háttérzajokat, szuszogásra nyakat ír ki.

Zajra nagyon érzékeny, nem léphez senki a szobába, köhögni, púzni nem lehet, széket nem lehet huzigálni.

Meg kell találni a megfelelő tempót és hanglejtést, valamint a zaj nélküli (kilégzés, belégzés) mikéntjét, utána határozottan jól működik. Szokni kell.

Egy elhangzott, vagy elhangzás nélkül beírt pont után már nem lehet tovább diktálni. Nem világos számomra, hogy honnan lehet tudni, hogy a program szerint vége van a mondatnak, vagy vár folytatást.

Adaptáció után olyan érzésem volt, hogy talán történt némi javulás a felismerésben.

Találati pontosság adaptácó előtt: 91.55% *Adaptáció után: 93.20%*

8.2.2. Szakértői vélemény

Az “Orvosi Íródeák” szoftver klinikai leletezésben való alkalmazhatóságáról

Prof. Dr. Pávics László
A Szegedi Tudományegyetem Orvostudományi Kara Nukleáris Medicina Intézetének vezetője

Az “Orvosi Íródeák” diktálóprogram tesztelésében intézetünk részéről öt fő pajzsmirigy-scintigramok leletezésében gyakorlott orvoskolléga vett részt, a Mesterséges Intelligencia Kutatócsoport szoftverfejlesztő munkatársaink iránymutatását követve. A szoftvert egy 1800 Mhz-es Pentium IV-es számítógépre telepítettük fel, a telepítésben és a mikrofon helyes beállításában a program készítői segítettek.

A program a legtöbb tesztelő személy hangjára már az első használatba vételkor jó hatékonysággal működött. Tapasztalatunk szerint azonban némi gyakorlást igényel annak a hanghordozásnak és beszédsebességnek a megtalálása, amely mellett a rendszerből a maximális felismerési hatékonyságot ki lehet hozni. A program működésének kiismerése után olyan találati pontosságot nyújtott, amelyet elégségesnek ítélek a gyakorlatban való alkalmazáshoz. Megítélésem szerint a program által kínált diktálási funkció segítségével a leletek számítógépbe való bevitele valóban kényelmesebbé és gyorsabbá tehető.

A tesztelést végző kollégáink legtöbbjének már a program alapbeállítási mellett is elfogadható találati pontosságú diktálást sikerült elérnie. A beszélőadaptációs lehetőség segítségével további javulást sikerült elérni a pontosságban, főként azon kolléga hangja esetén, akire a rendszer alapállapotban kicsit gyengébben működött. A legtöbb gondot az okozta a tesztelés során, hogy a rendszer nagyon érzékeny a háttérzajokra. Ez a probléma azonban megfelelő mikrofon segítségével valószínűleg kiküszöbölhető.

A program által támogatott szókincs és nyelvi szerkezetek összhangban állnak az intézetünknél pajzsmirigy leletek szövegezésében használatos nyelvezettel. Így a program a

jelenlegi nyelvi beállítások mellett lehetővé teszi, hogy a szokványos felépítésű pajzsmirigy-szcintigram leleteket diktáljunk. Remélhetőleg a nyelvi rész későbbi kibővítésével és lecserélésével a program alkalmazhatóvá válik majd más tematikájú leletek diktálására is.

A program felhasználói felülete egyszerű, könnyen kezelhető. A tesztelés során a program kezelésével kapcsolatban lényeges problémánk nem adódott. Azonban mivel a program segítségével készített jelentések szövege a felhasználói felület jelen formája mellett nem, vagy csak nehézkesen vihető át a tanszékünkön szokványosan alkalmazott leletnyilvántartó programba, ezért a későbbiekben mindenképpen érdemes lenne a diktálóprogramot integrálni szokványos orvosi adatkezelő és archiváló rendszerbe.

Dr. Pávics László