

Tanulmány

A GENOMIKUS MEDICINA INFORMATIKÁJA

Deutsch Tibor

a kémiai tudomány kandidátusa, tudományos
főmunkatárs, Semmelweis Egyetem,
Alkalmazott Logikai Laboratórium

Gergely Tamás

a matematikai tudomány doktora, igazgató
Alkalmazott Logikai Laboratórium

Bevezetés

Az egészségügyben tapasztalható romló hatékonyság és költségrobbanás világszerte gondokat okoz, amit csak tetéz, hogy az ellátás minősége nem növekszik a ráfordításokkal arányosan. A jelentkező problémák egyik meghatározó oka az, hogy nem tudjuk hatékonyan megelőzni a krónikus betegségek kialakulását, melyek menedzselése jelentős erőforrásokat emészt fel. A krónikus beteg ugyanis tartós ellátást (főként gyógyszerelést), jelentős személyi ráfordítást (nővérórák, orvosi törődést) igényel, továbbá terheli a társadalombiztosítást, és (ha munkaképes korban van) alkalmanként kiesik a gazdasági tevékenységből is. A költségeket tovább növeli, hogy a jelenlegi megelőző és szűrőprogramok nem kellőképpen veszik figyelembe az egyéni kockázati tényezőket, és a diagnosztikus, valamint terápiás tervek sem minden esetben alkalmazkodnak a betegek egyéni sajátosságaihoz. A konfekcionált gyógyítást mozdítják elő a különböző szakmai irányelvek és protokollok is, noha ezek korántsem kínálnak optimális ellátást valamennyi esetben.

A jelenlegi orvosi gyakorlat ezért szükségképpen reaktív, azaz rendszerint a már bekövetkezett eseményekre (például tünetek, panaszok megjelenése) reagál. A terápiá

iteratív titrálása átmeneti kudarcokon keresztül halad előre, ami természetesen megterheli a beteget, és szükségtelen költségekkel jár. Természetesen más volna a helyzet, ha egyes genetikai eltérésekről tudnánk, hogy elősegítik bizonyos betegségek kialakulását, mert ilyenkor megfelelő intézkedésekkel megelőzhetnénk (de legalábbis késleltethetnénk) a kórfolyamatokat. Ez nemcsak a beteg és orvos szempontjából volna előnyös, hanem jelentős gazdasági haszonnal is járna, hiszen a széleskörű megelőzésre fordítandó pénz sokkal kisebb annál az ösz-szegnél, amelyet a krónikus betegségek terápiájára kell költeni.

Az előrelátáson alapuló (proaktív) és egyénre szabott ellátás szakmai feltételeit hosszabb távon elsősorban az Emberi Genom Projekt és az ehhez kapcsolódó molekuláris biológiai kutatások eredményei hivatottak megteremteni. Az orvoslásnak ezt az újágát nevezik *genomikus* (genetikai ismeretekre alapozott) *medicinának*, melynek lényege, hogy az orvos az egyéni genetikai eltérésekhez alkalmazkodva tervezi meg és alkalmazza a különböző megelőző, szűrő és gyógyító eljárásokat (Guttmacher – Collins, 2002).

Míndez azonban csak oly módon képzelhető el, hogy a különböző orvosi és bioinformatikai módszerek segítségével a genomikai

és orvosi ismereteket kombinálni tudjuk a páciensek adataival (Sarachan et al., 2003). Az érintett tudományterületek kapcsolódását és átfedését az 1. ábrán szemléltetjük.

Az orvosi informatika és bioinformatika összeolvadásából létrejövő új tudományág a biomedicinális informatika. Ebben a tanulmányban ezt szeretnénk bemutatni az olvasóknak. Mielőtt azonban az informatikai módszerekre rátéménk, röviden bemutatjuk a genomikus orvoslást, melyből kirajzolódik, hogy annak gyakorlati megvalósításához milyen számítógépes technológiákra van szükség.

Genomikus medicina

A genomikus medicina genetikai ismereteken nyugvó, előrelátáson alapuló, egyénre szabott megelőzést és gyógyítást jelent. Lényege, hogy (1) az egyéni genetikai eltérések alapján igyekszünk előrelátni a fenyegető kórfolyamatokat, (2) mindent megteszünk ezek megelőzéséért és/vagy késleltetéséért, (3) molekuláris diagnosztikai eszközökkel a lehető legkorábban igyekszünk észlelni

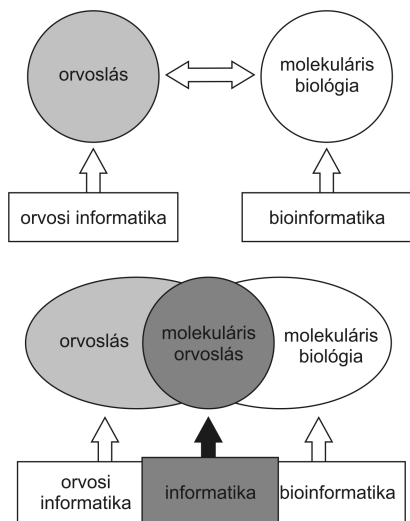
a kezdődő elváltozásokat, és (4) a páciens egyéni sajátosságaihoz alkalmazkodva olyan beavatkozásokat alkalmazunk, melyek lehetőség szerint normalizálják/stabilizálják a beteg állapotát, és csökkentik a további kockázatokat.

Az orvoslás ezen új ágának a fentiek értelmében feltétele, hogy ismerjük a betegségek és terápiás beavatkozások molekuláris mechanizmusait, és azt is, hogy az emberi genomon talált eltérések miként befolyásolják a különböző betegségek kialakulását. Ezen általános ismeretek mellett természetesen arra is szükség van, hogy meghatározzuk az adott páciensben jelentkező genetikai eltéréseket, kórelőzményeket, továbbá a beteg általános egészségi állapotát, környezetét és életmódját.

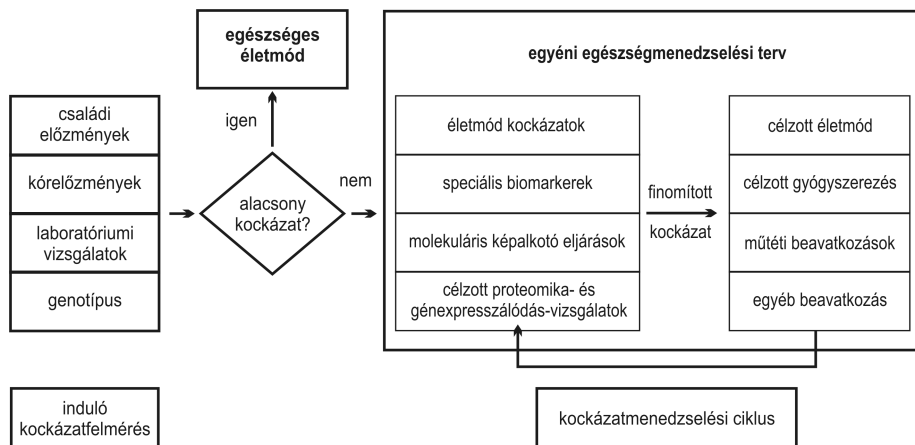
Ami a genetikai háttérrel illeti, számos betegség (például hipertónia, diabétesz és ischaemiás szívbetegségek) kialakulásában több gén is szerepet játszhat, így a kóros eltérések lehetséges variációja is meglehetősen nagy. Ezért nem elég pusztán a betegségek megjelenési formáját (fenotípusos manifestáció) vizsgálni, hanem a pontos diagnózishoz a genotípust, tehát az adott egyéni genetikai jellemzőit, a betegség kialakulásának folyamatát (etiológia), és az ennek során lejátszódó sejten belüli (subcelluláris) folyamatokat is ismerni kell. A kóros felnőttkori elhízások 1-6 %-ában például a melanocortin-4 receptor mutációja játszik szerepet, és ezeket az eseteket meg kell különböztetni a kóros elhízás egyéb típusaitól.

Ezt a fajta sokrétű vizsgáldást nemcsak a diagnózis felállítása, hanem a terápia megtervezése esetében is el kell végezni, hiszen a genomika eredményei azt is lehetővé teszik, hogy a hagyományos terápiás módszereket egyénre szabottan alkalmazzuk, sőt teljesen újszerű módon avatkozzunk be a kórfolyamatokba.

Az emberi genomtól a genomikus orvoslásig vezető lépéseket a 2. ábra mutatja.



1. ábra • Genomikus medicina és biomedicinális informatika



2. ábra • A kockázatmenedzselés folyamata

Genetikai elváltozások feltérképezése

A genomikus medicina fontos eleme a lehetséges genetikai elváltozások feltérképezése. Korszerű vizsgáló módszerekkel ma már kis mennyiségű vérből ki tudjuk mutatni azokat a genetikai mutációkat és polimorfizmusokat, melyek növelik bizonyos megbetegedések valószínűségét. Noha ezek a vizsgálatok jelenleg még speciális laboratóriumokban történnek, a közeli jövőben megjelenő biochipek segítségével már szélesebb körben is el fogjuk tudni végezni ezeket a tesztekkel, és a beépített számítógépes programokkal gyorsan ki is tudjuk értékelni a kapott adatokat.

A reális megítélés érdekében azért érdemes megjegyezni, hogy ennek a módszernek – legalábbis egyelőre – még többféle

(szakmai, anyagi és etikai) korlátja létezik. Szakmai ismereteink hiányosságai nyilvánvalóak, hiszen többnyire még nem ismerjük pontosan azokat a lehetséges genetikai variációkat, amelyek egy adott kórkép kialakulásában szerepet játszhatnak, és azt sem tudjuk, hogy ezek az eltérések egymáshoz képest milyen szerepet játszanak. A molekuláris mechanizmusok felderítése terén is sok még a tennivaló. Bizonyos betegségek esetén már valószínűleg lehetséges az itt leírt eljárások bevezetése, de az általános alkalmazástól még messze vagyunk.

Az anyagi vonatkozások tulajdonképpen alig igényelnek kifejtést, hiszen a molekuláris genomikai kutatás igen drága (ami korlátozza az egyszerre vizsgálható betegségek körét), és jelenleg a szükséges diagnosztikai eszközök

Kockázat fokozata

- Alacsony
- Magas

Krónikus betegség (korai fázis)

Krónikus betegség (kései fázis)

Egészségmenedzselési terv

- generikus egészségmegőrzési terv (*wellness pathways*)
- kockázat csökkentése egyénre szabott életmóddal, terápiás és szűrési programmal
- rövid és hosszú távú kockázatok csökkentése egyénre szabott terápiával és monitorozással
- életminőséget javító és palliatív módszerek

1. táblázat • Kockázatok és menedzselési módok

zők is meglehetősen költségesek. Várható, hogy a technológia fejlődésével a költségek csökkennek majd, de a mindennapi használat egyelőre elég távolinak tűnik.

Noha a lehetséges etikai problémák zömét még nem is ismerjük, a már felmerült etikai kérdések is nagy vitát váltottak ki. Csak néhány mutatóba anélkül, hogy ebben a cikkben ezekről részletesen szólnánk: Van-e jogunk egy beteg genetikai anyagát feltérképezni; ki és milyen feltételekkel férhet hozzá ezekhez az adatokhoz; milyen korlátokat – esetleg előnyöket – jelenthet a genetikai diszpozíció ismerete; mikor és kinek a kezdeményezésére és hozzájárulásával kell, illetve lehet beavatkozni, stb.? Mindenképpen célszerű, ha az olvasók ezekre a dilemmákra is gondolnak, miközben megismerkednek a genomikus orvoslás informatikai módszereivel.

Kockázatmenedzselés

Az elvégzett genomikai és klinikai vizsgálatok alapján egy adott személyt az 1. táblázatban szereplő négy kockázati csoport valamelyikébe sorolhatjuk be. A különböző csoportokhoz természetesen eltérő ellátási módszerek is tartoznak. Természetesen nem mindegy, hogy a vizsgált személy életének mely szakaszában kerül vizsgálatra, hiszen egy gyerek esetében egészen másként kell az adatokat értékelni, mint egy idősebb beteg esetében. A gyerekkorban felismert kockázat egész életre szóló programot jelent, míg egy adott elváltozás kockázatának késői felismerése elsősorban a következmények gyors mérséklését teheti indokolttá.

Alacsony kockázatú személyeknél a fő feladat a kockázati tényezők keletkezésének megakadályozása és az egészség megőrzése. Ilyenkor a megelőzési terv kidolgozásánál jórészt általános irányelvekre támaszkodhatunk.

Felmerülhet persze, hogy miért kerül egy alacsony kockázattal rendelkező személy egyáltalán vizsgálatra. Erre ma még nehéz

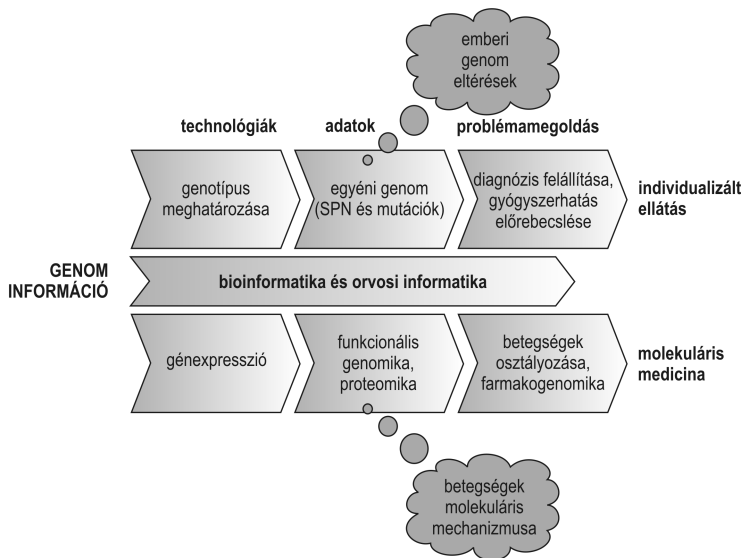
pontosan válaszolni. Bizonyos – például gyakori vagy veszélyes – betegségcsoportok esetén elképzelhetők általános szűrővizsgálatok (ahogy már ma is vannak ilyenek, bár más módszerekkel); lehetséges, hogy környezeti vagy történeti tényezők teszik indokolttá a vizsgálatot (például családi halmozódás, *malignus* környezeti tényezők stb.); végül bizonyos foglalkozások gyakorlásához előírhatnak általános orvosi vizsgálatot is (ahogy ez ma is történik, bár ez egyelőre genetikai adatokra nem terjed ki). Az azonban – legalábbis egyelőre – a fantázia világába tartozik, hogy minden ember esetében például születéskorteljes genetikai térképet készítsenek, bár elvileg ez sem lehetetlen (lásd a fentebb elemzett korlátokat!).

Magas kockázat esetén ezzel szemben egyre speciálisabb, precízebb és költségesebb eljárásokkal kell pontosan tisztáznunk a kockázati profilt. A kockázatok természetesen időben is változhatnak (változik például a környezet, az egyéni életvitel, és/vagy maga a genetikai anyag a késői gének aktivizálódása vagy mutációk révén), és ezért a becsléseket időszakosan fel kell frissíteni.

A kockázatbecslés egy időben zajló, többlépéses, ciklikus folyamat, melynek sémáját a 3. ábra szemlélteti.

A kockázatbecslés végső célja természetesen egy *egészségmenedzselési terv* kialakítása, amely (a krónikus betegségek végső fázisától eltekintve) a kockázatok csökkentését és a kórfolyamatok korai észlelését célzó, egyéni szabott beavatkozásokból épül fel.

Az elsődleges és másodlagos megelőzés (prevenció) esetében nyilvánvalóan *kockázatcsökkentés*ről van szó, de a krónikus betegségek ellátása is lényegében ezt jelenti. A beavatkozások intenzitása ugyanis elsősorban attól függ, hogy a megismert adatok mekkora valódi kockázatot jelentenek. Frissen felfedezett cukorbetegségben például a magas vércukorszint főként nem önmagában veszélyes, hanem azért, mert hosszú távon



3. ábra • A genomtól a genomikus medicináig

súlyos szövődeményeket okozhat. Ezért más-ként kell kezelni egy újszülöttkori, egy fiatal felnőttben jelentkező vagy éppen egy időskori cukorbetegséget, noha a tünetek gyakran azonosak vagy hasonlóak lehetnek.

A kockázatok figyelembevételével határozhatjuk meg a célszerű szűrési módszereket és ezek gyakoriságát is. A National Center for Toxicogenomics által kifejlesztett ToxChip segítségével olyan genetikai eltéréseket mutathatunk ki, melyek a környezeti ártalmak hatását befolyásolják.

A kockázatmenedzselés fontos része a kórfolyamat korai észlelése. Ma már léteznek molekuláris képalkotó eljárások, amelyekkel láthatóvá tehetjük a sejtekben kibontakozó kórfolyamatokat. Hasonló célt szolgálnak a különböző biomarkerek is, azaz az egyes folyamatok során időben változó mennyiségben megjelenő, jellemző és kimutatható molekulák, melyek jelenléte már a kórfolyamat kezdetén jelezheti az elváltozásokat.

A kialakult elváltozások/betegségek gyógyítását szintén az egyéni sajátosságokhoz kell illeszteni. Ha egy adott páciens esetében tudjuk,

hogy a kóros elhízás a melanocortin-4 receptor mutációjának következménye, akkor olyan készítményt célszerű alkalmazni, mely a szóban forgó receptorhoz kötődik. Amennyiben a cytochrome P450 mitokondriális enzim kódoló génben észlelünk olyan polimorfizmust, mely jelentősen csökkenti a máj metabolizációs képességét, kisebb adagban kell alkalmazni mindazokat a készítményeket, melyek főként a májon keresztül ürülnek ki a szervezetből.

Megjegyezzük, hogy a krónikus betegségek végső fázisában már kisebb szerep jut a genetikai eltéréseknek, hiszen ilyenkor már az életminőség javítása és a szenvedés enyhítése kerül előtérbe.

Biomedicinális informatika

A 2. ábrán jól látható, hogy a genomikus orvoslásban fontos szerepet játszanak a különböző bioinformatikai és orvosi informatikai módszerek. A bioinformatika olyan eljárásokat és modelleket kínál, melyek segítségével betekinthetünk a betegségek molekuláris mechanizmusaiába, és megért-

hetjük a betegségek lefutásában, illetve a terápiás hatásokban megfigyelhető egyéni eltérések okait. Az orvosi informatika esz-köztárában olyan módszerek szerepelnek, melyek biztosítják a betegekkel kapcsolatos információk hatékony szervezését, és segítik az orvosi döntések meghozatalát (Deutsch – Gergely, 2003). A két szakterületről származó technikák integrálásával olyan informatikai szolgáltatásokhoz jutunk, melyek hozzájárulnak a kockázatok felméréséhez, s ezek alapján az optimális tennivalók meghatározásához.

A biomedicinális informatika módszereit többféle szempont szerint csoportosíthatjuk. A 2. táblázatban a biológiai szerveződés különböző szintjeihez kapcsolódóan soroltuk fel a legfontosabb alkalmazásokat.

Terjedelmi okok miatt a továbbiakban csupán a különböző adatbázisokkal, számítógépes szimulátorokkal és döntéstámogató konzulensekkel foglalkozunk.

Adatbázisok

A klinikus alapvető igénye, hogy könnyen

hozzáférjen a legújabb tudományos eredményekhez és a betegek adataihoz, továbbá segítséget kapjon azok feldolgozásához és elemzéséhez.

Tudományos eredmények és adatok

A GenBank (Genetic Sequence Database), a SWISS-PROT (Protein Sequence Knowledge Base), és az IMGT (Immunogenetics Database) adatbázisok különböző gének és fehérjék szerkezetének leírását tartalmazzák. Ezek az adatbázisok ingyenesen hozzáférhetők az Interneten keresztül, és standard adatmodelljüknek köszönhetően azonos elvek szerint lehet adatot keresni bennük. A különböző gének expresszáldására vonatkozó adatokat ezzel szemben már eltérő sémák szerint adják meg.

A KEGG (Kyoto Encyclopedia of Genetics and Genomics), Bind (The Biomolecular Interaction Network Database), Pronet (Protein Interaction Network Database) és Transfac (The Transcription Factor Database) adatbázisokban különböző gének és fehérjék bonyolult kölcsönhatásai szerepelnek (Kane-

Információs szint	Informatikai eszközök
Molekuláris	<ul style="list-style-type: none"> · genotípus automatizált meghatározása · genomikai adatbázisok
Sejt	<ul style="list-style-type: none"> · számítógéppel segített molekuláris képkalkotó módszerek · DNS-chip vizsgálati adatok feldolgozása és értelmezése · funkcionális genomikai, proteomikai és metabolomikai adatbázisok · sejtszintű szabályozási és funkcionális modellek
Szövet/szerv adatok (adatok)	<ul style="list-style-type: none"> · integrált hisztológiai és genomikai adatbázisok (pl. tumor-adatok) · strukturális és funkcionális patológiai modellek
Emberi szervezet/beteg	<ul style="list-style-type: none"> · genetikai adatokkal kibővített egészségi/kórlap · komplex állapotmodellek · betegségek újracsatározása molekuláris mechanizmusok alapján · farmakogenomikai adatbázisok · in silico betegségmodellek · genetikai alapú megelőzési és klinikai irányelvek
populáció	<ul style="list-style-type: none"> · genomikus epidemiológiai adatbázisok és számítógépes hálózatok · genetikai adatokkal kibővített <i>outcome</i> adatbázisok

2. táblázat • Informatikai módszerek alkalmazása különböző szerveződési szinteken

siha – Goto, 2000). A PubGene szoftver ezeknek a kölcsönhatásoknak a feltérképezését segíti.

Ha egy adott páciensben az orvos valamilyen azonosított genetikai eltérést talál (lásd fentebb), akkor a felsorolt (és további hasonló) adatbázisokban kereshet adatokat arra vonatkozóan, hogy ennek az eltérésnek milyen molekuláris és biokémiai változások lehetnek a következményei.

Az orvosi gyakorlat számára különösen fontosak azok az adatbázisok, melyekben azt találhatjuk meg, hogy a különböző genetikai eltérések milyen morfológiai és funkcionális változásokat okoznak, és miként függnek össze a különböző betegségek kialakulásával. A tumor-adatbázisok például azt tartalmazzák, hogy a különböző polimorfizmusok és mutációk hogyan befolyásolják a daganatok növekedését. Az OMIM (Online Mendelian Inheritance in Man) adatbázis jelenleg több mint kilencezer olyan gént tartalmaz, melyek működése összefügg valamilyen betegséggel.

Nagyon hasznosak azok az információforrások is, melyekben a különböző genetikai eltéréseknek a különböző gyógyszerek hatásaira gyakorolt befolyására találunk adatokat (Roses, 2002). Ezek mintapéldája a Stanford Egyetemen készült PharmKKB (The Pharmacogenetics and Pharmacogenomics Knowledge Base) ismerettár.

A fentiekhez hasonló adatbázisok létrehozása és folyamatos karbantartása korántsem egyszerű feladat. A rohamos tempóban bővülő szakirodalomból ugyanis egyre nehezebb kigyűjteni a legfrissebb eredményeket, és ezeket naprakészen és megfelelő formában beilleszteni a megfelelő adatbázisokba. A közeljövőben egyre inkább rászorulunk majd olyan információkinyerő/feldolgozó robotok segítségére, melyek rendszeresen figyelik és meg is értik a megjelenő szakirodalmat, és arra is képesek, hogy folyamatosan karbantartsák a

különböző adatbázisok tartalmát. Ezek az intelligens eszközök arra is használhatók lesznek, hogy az új eredményeket a már tárolt ismeretek alapján értelmezzék, és felhívják a figyelmet az esetleges ellentmondásokra. A robotok fejlettebb változataitól azt is elvárhatjuk, hogy a közleményekben szereplő ábrákat és grafikonokat is megértsék, vagy egy szöveges közleményből automatikusan megszerkesszék a vizsgált fehérjék kölcsönhatásának térképét.

Azért mindez továbbra sem nélkülözheti az emberi felügyeletet, noha kétségtelen, hogy ekkora adatmennyiséget az ember már nem képes áttekinteni. Az úgynevezett szakértői rendszerek fejlesztése – mint amilyenek az itt említettek is – ezért az emberi ismeretek és a gépi feldolgozás állandó kölcsönhatását igénylik, és még igen fejlett változataik sem nélkülözhetik időnként az emberi tapasztalatot és intuíciót.

Az intelligens adatfeldolgozó rendszerek fejlesztésének kedvez a szakmában egyre szélesebb körben terjedő standard terminológia, az egységes – stilisztikailag éppenséggel nem szép, de nagyon praktikus – szóhasználat, amelynek rugzítása és elterjesztése a fejlesztés egyik kulcskérdése. Érdeemes megjegyezni még azt is, hogy az ilyen általánosan használható robotrendszer fejlesztése még a kezdeteknél tart, vagyis még jelentős időnek kell eltelnie ahhoz, hogy ezeket az intelligens eszközöket széles körben alkalmazni lehessen.

Orvosi adatbázisok

A molekuláris biológiai adatbázisoktól eltérően a különböző egészségügyi kórlapokon tárolt adatok mindig egy adott páciensre vonatkoznak, és optimális esetben valamilyen információt tartalmaznak, ami orvosi szempontból fontos lehet. A genomikus medicina gyakorlatában használt egészségügyi rekordokban azonban a megszokott adatokon kívül természetesen tárolni kell

az adott személy genomikai adatait, és mindazokat az információkat (környezeti ártalmak, életmód stb.) is, melyek szerepet játszanak a megelőzésben és gyógyításban (Martin-Sanchez et al., 2002).

Nemcsak a rögzített adatok köre, hanem az új kórlap felépítése is eltér a megszokottól. Az új kórlapon a páciens állapotának leírására az ún. *egészségprofil*t használjuk, mely a szubjektív és objektív (mért) adatok összefoglalására szolgál. Az állapot megítélésénél a megszokott diagnózisokhoz és problémákhoz társul a páciens komplex állapotleírása, és ebben a részben kapnak helyet a különböző genetikai, környezeti és életmódbeli kockázatokra vonatkozó bejegyzések is. Az egészségügyi ellátási terv elemei között a tervezett vizsgálatokat és terápiás beavatkozásokat, valamint azok időbeli ütemezését is megtaláljuk.

A kórlapok gyakorlati felhasználásához az is szükséges, hogy azok tartalmazzák a legkülönbözőbb ellátási eseményeket függetlenül attól, hogy ezek hol (kórházban, orvosi rendelésben vagy a beteg otthonában) történtek. Ezeket a kórlap-adatbázisokat mindig a gyakorló orvosoknak, illetve az ápolószemélyzetnek kell karbantartania, és ott kell őket tárolni (vagy legalábbis hozzáférhetővé tenni), ahol a beteg kezelése vagy nyomon követése folyik. Az adatokhoz történő hozzáférést ugyanakkor szigorú szabályokkal kell korlátozni, és kiemelt figyelmet kell fordítani a genetikai adatok gyűjtésére, tárolására és feldolgozására.

Az egyéni egészségügyi rekordok és kórlapok adataiból nagyméretű adatbázisokat hozhatunk létre. Ezek prototípusa az Egyesült Államokban működő HuGENet (Human Genome Epidemiology Network) hálózat, de számos hasonló projekt fut Európa különböző országaiban is. A jelenleg érvényes adatkezelési szabályoknak megfelelően a különböző epidemiológiai adatbázisok elsősorban statisztikai jellegűek. Ezekben

az összegző adatbázisokban különböző hozzáférési jogosultságokat határoznak meg, például az egyes betegek adataihoz csak a kezelőszemélyzet és néhány erre felhatalmazott vezető férhet hozzá, míg az összesítő adatok jóval szélesebb körben (esetenként mindenki számára) hozzáférhetők.

Miként férünk hozzá a szükséges adatokhoz?

A különböző adatbázisok gyakorlati használata természetesen attól függ, hogy milyen könnyen és pontosan tudjuk megtalálni bennük a szükséges információt. A genomikus medicina esetében a helyzetet bonyolítja, hogy a különböző adatbázisok között sok az átfedés, és a klinikus nem mindig tudja, hogy a kért adatot melyikben találhatja meg. A lekérdezéshez tehát olyan keresőprogramokra van szükség, melyek ezekből a heterogén és elosztott információtarakból is képesek kikeresni a kért információt.

Nem szabad elfelejteni azt sem, hogy a klinikusok többsége informatikai szempontból kevésbé képzett (tisztelt a kivételnek!). Ezért olyan adatkereső rendszereket kell kifejleszteni, amelyek lehetőleg kevés előképzettséggel is érdemi információgyűjtést biztosítanak. A klinikus számára lehetővé kell tenni, hogy kulcsszavakkal, szövegesen és grafikusán egyaránt megfogalmazhassa igényeit. Ehhez olyan keresőeszközökre is szükség van, melyek megértik a felhasználó szövegesen megadott kéréseit, és képesek megtalálni például egy biokémiai térkép bizonyos részletét is. A lekérdező programokhoz olyan adatmegjelenítő eszközöket kell kapcsolni, melyek a felhasználók igényei szerint jelenítik meg a találatokat.

Mindez persze mindaddig csak álom, ameddig a különböző adatbázisokat nem tudjuk összekapcsolni egymással. Az integrálásnak számos akadálya van. Már az is gondot jelent, hogy a hasonló célú klinikai és genomikai adatbázisok sem kompatibi-

lisek egymással. Az egyes ellátóhelyeken különböző adatokat tárolnak, és különböző fogalomtárakat és szótárakat (például UMLS (Unified Medical Language System), SNO-MED (Systematised Nomenclature in Medicine), MeSH (Medical Subject Headings), és ICD (International Classification of Diseases) használnak. A nemzetközi szabványok, például a HL7 (Health Level Seven) protokoll gyakorlati alkalmazása a tapasztalatok szerint számos nehézséggel ütközik.

Ugyanezzel a sokféleséggel találkozunk a molekuláris biológiában használt adatbázisok körében is. Az MGED (Microarray Experiments DataBase) adatbázis például a génexpresszálódási vizsgálatok adatainak leírására szolgál. A TAMBIS (Transparent Access to Multiple Bioinformatics Information Sources) szótár a bioinformatika és molekuláris biológia területén használt fogalmakat fedi le, de ettől lényegesen eltérő struktúrájú szótárakat hozott létre például a Gene Ontology Consortium is.

A nehézségeket tovább fokozza, hogy nincsenek standard szótáraink a fenotípus-, környezeti és életmódjellemezők, fejlődési stádiumok, patológiák és toxikológiai fogalmak egységes megnevezésére sem. A legfrissebb ICD és CPT (Current Procedural Terminology) szótárakban nem vagy csak töredékesen szerepelnek genetikai fogalmak és a genomikus medicinában használt diagnosztikus, valamint terápiás eljárások, továbbá ezek leírása sem mindig pontos és egyértelmű.

A múltbeli tapasztalatok alapján és az ismeretek ugrásszerű növekedését látva sajnos abban sem igen reménykedhetünk, hogy a közeljövőben majd létrejön egy átfogó és standard orvosi fogalomtár. A feltétlenül szükséges integrálást ezért csak olyan eszközökkel képzelhetjük el, melyek lehetővé teszik az átjárást a létező különböző fogalomtárak és kódrendszerek között. Ezt segítik elő azok a referencia-ontológiák, melyek a legkülönbözőbb adatokat és ismereteket kü-

lönböző szerveződési szinteken (sejt, szövet, szerv, szervezet és populáció) rendszerezik és tárolják, ugyanakkor képesek arra, hogy egységesen használják a különféle szinteken használt fogalmakat.

Érdemes megjegyezni, hogy az ilyen átfogó és többszintű rendszerek készítésének nem elvi, sőt nem is gyakorlati akadályai vannak, hiszen a szükséges matematikai és informatikai eszköztár már ma is rendelkezésre áll. Az akadály sokkal inkább anyagi, pontosabban profittermészetű, a készítők versengésében a belátásnak kevés a szerepe, miközben sokféle presztízsszempontra is érvényesül.

Számítógépes szimulátorok

A klinikus számára nagyon fontos, hogy kellő megbízhatósággal előre meg tudja becsülni a betegségek kialakulását és a tervezett beavatkozások várható hatásait. A jelenleg rendelkezésre álló nagyszámú morbiditási és túlélési összefüggések is ezt a célt szolgálják. Egyes képletekkel például egy bizonyos időtartományban kellő megbízhatósággal ki tudjuk számítani a szívizominfarktus valószínűségét, ha ismerjük a beteg nemét, életkorát, szérumkoleszterin-szintjét és dohányzási szokásait. Ezekkel az előrejelzésekkel azonban rendszerint az a probléma, hogy csupán bizonyos demográfiai, klinikai és kockázati tényezőket, illetve azok kombinációját tartalmazzák, és a fenti esetben például már semmit sem mondanak arról, hogy ez a kockázat miként alakul cukor- vagy hipertóniás betegeknél, vagy olyanoknál, akik valamilyen koleszterinszint-csökkentő gyógyszert szednek. Ám témánk szempontjából még fontosabb, hogy a hagyományos epidemiológiai összefüggésekben egyáltalán nem szerepelnek genetikai paraméterek.

A megfelelő előrelátáshoz olyan számítógépes (*in silico*) betegségmodellekre van szükség, melyek egy adott *T* időpontban elvégzett állapotfelmérésből, továbbá a beteg

genetikai, demográfiai, klinikai és életmódbeli jellemzőiből következtetni tudnak a t idő elteltével várható fejleményekre. Az általános M betegségmodellt a következő összefüggéssel fogalmazhatjuk meg:

$$BP(T + t) = M[BP(T), \acute{A}P(T), GE, KLP, KP, \acute{E}M, TH]$$

A fenti modellben a betegség progresszióját valamilyen jellemző paraméter (BP) időbeli változásával jellemezzük (cukorbeteg esetében BP lehet például az éhomi vércukor értéke, mely kezeletlen betegeknél fokozatosan emelkedik a betegség előrehaladásával). A betegség progresszióját a beteg általános egészségi állapota (ÁP), az észlelt genetikai mutációk és polimorfizmusok (GE), a beteg demográfiai és klinikai paraméterei (KLP, például nem, életkor, inzulin-érzékenység stb.), a környezeti paraméterek (KP, például légszennyezettség mértéke, sugárterhelés nagysága stb.), a beteg életmódja (ÉM, például fizikai aktivitás mértéke, étkezési szokások, dohányzás stb.) és az alkalmazott terápia (TH, például gyógyszerelés) határozza meg.

Egy adott betegségre (például diabétesz) kidolgozott modell rendszerint nagyszámú (több száz) matematikai összefüggést tartalmaz, és az előrelátó (proaktív) medicina sikere természetesen attól függ, hogy ezek segítségével mennyire megbízhatóan és milyen időtávlatban tudjuk előre látni a betegség kialakulását és progresszióját. A modell viselkedését számítógépes szimulációval tanulmányozhatjuk, különböző változatokat játszhatunk le, miközben egy vagy több feltételt változtatunk.

A fenti *in silico* betegségmodelleket alapvetően kétféle módon hozhatjuk létre. A célszerű módszer kiválasztása a probléma természetétől függ.

Bottom-up modellezés

A *bottom-up* modellek a rendszerbiológia jellegzetes eszközei, melyekkel az alacso-

nyabb szinteken zajló folyamatok modelljeiből építőkocka-szerűen konstruálják meg a magasabb szintű életfolyamatok működési modelljét. Az egyes sejtekben zajló folyamatok modellezéséhez például meg kell fogalmaznunk a különböző gének és fehérjék közti kölcsönhatásokat és az anyagcsere bonyolult biokémiai folyamatait. A sejten belüli folyamatok leírására leggyakrabban különböző hálózati modelleket alkalmazunk (Jeong et al., 2000).

10-15 éven belül várhatóan létre fogjuk hozni egy sejt teljes, működő számítógépes modelljét. Erre irányulnak a jelenleg futó E-Cell és Virtual Cell projektek (Tomita, 2001). A résztvevők együttműködését egységes adatformátumok és modellezési nyelvek (SBML (Systems Biology Markup Language) és CellML (Cellular Markup Language) teszik lehetővé. Az E-Cell projekt célja például egy olyan virtuális emberi vörösvérsejt létrehozása, mely minimális génállománnyal (127 gén) rendelkezik.

A sejtekből felépülő szövetek és szervek morfológiájának és funkcióinak leírására főként olyan nem-lineáris dinamikussal modelleket alkalmaznak, melyek képesek megragadni a sejtstruktúrák önszerveződési folyamatait. Az Oxfordi Egyetemen nagyszámú virtuális szívizomsejtből virtuális szívet hoztak létre, melynek segítségével vizsgálni tudták, hogy a különböző patológiás állapotok milyen eltéréseket hoznak létre a szív elektromos aktivitásában, vagy a beteg szív miként reagál különböző gyógyszerekre (Noble, 2002).

A *bottom-up* modellépítés végső célja természetesen a teljes emberi szervezet valóság-hű modelljének létrehozása. Ebbe a törekvésbe illeszkedik a Virtual Patient és Physiome projekt. Az Entelos cég munkatársai többek között elkészítették az asztmás beteg modelljét, melyben genetikai és patofiziológiai paraméterek mellett életmódtényezők is helyet kaptak. A mintegy 3500 tudományos közlemény alapján felállított egyenletekben közel

7500 paraméter szerepel, melyek az asztmás rohamok kialakulásában szerepet játszó tényezők hatását tükrözik. A különböző paraméterek értékeit egyedileg lehet beállítani, s ezáltal azt is tanulmányozni lehet, hogy a különböző gyógyszerek különböző páciensekben miként enyhítik az asztmás rohamokat (Trimmer, 2001).

A Physiome projektben szintén építő-kockaszerűen kombinálják a génekre, fehérjékre, sejtekre, szervekre és az anyagcsere-folyamatokra vonatkozó modellrészeket (Bassingthwaight et al., 2002). A modell-építés sémáját a 4. ábrán mutatjuk be.

A minden tekintetben valósághű virtuális páciens létrehozásától még meglehetősen távol vagyunk. A nehézségek sokrétűek. A különböző szintű ismereteket önmagukban is nehéz megfogalmazni egységes struktúrájú modellek formájában. Ennél is nagyobb szakmai nehézséget jelent azonban, amikor az alacsonyabb szinten megfogalmazott leírásokat egy magasabb szerveződési szintre akarjuk átvinni. Az emberi szervezet ugyanis nem egy mechanikus gép, mely különböző bonyolultságú elemek összeszerelésével készül, hanem egy olyan komplex, adaptív és

önszerveződő rendszer, melynek stabilitását és plaszticitását éppen a különböző szintek közötti bonyolult információs kapcsolatok biztosítják, és a betegségeket nem utolsósorban ezek sérülései idézik elő.

Top-down modellek

Jelenlegi ismereteink és a rendelkezésre álló számítási kapacitás mellett célszerűbbnek tűnik a betegségek *top-down* modellezése, amikor a klinikai manifesztációkból indulunk ki, és csak olyan mértékben használunk élettani, biokémiai és genomikai ismereteket, hogy figyelembe tudjuk venni a betegek azon egyéni jellegzetességeit, amelyek befolyásolják a betegségek progresszióját és a terápiák hatásait. Ennek kitűnő példája az Archimedes-modell, mely egy olyan ok-okozati és hatástérkép, melynek csomópontjaiban különböző patofiziológias állapotok, funkciók, panaszok, tünetek és kóros laboratóriumi értékek szerepelnek. A különböző csomópontokat összekötő élek különböző típusú kapcsolatokat fejeznek ki a csomópontok között (Schlessinger – Eddy, 2003). A különböző csomópontokhoz részletes leírás tartozik, míg az élekhez esetenként bonyolult matematikai kifejezések kapcsolódnak. A koronáriabetegség progresszióját leíró egyenletek például azt fogalmazzák meg, hogy hol és milyen mértékben szűkülnek az érpályák, és ezek nyomán mikor jelentkeznek az anginás panaszok. A matematikai összefüggések azt is tartalmazzák, hogy a csomópontokkal ábrázolt állapotok miként függnek a betegek egyéni tulajdonságaitól. A diabéteszmodellben például az inzulinérzékenység és a glükóztolerancia betegenként nyilvánvalóan eltér egymástól. Ha ezeknek az eltéréseknek genetikai okuk van, a modellben ezek a releváns polimorfizmusok és mutációk is megjelennek, és a megfelelő csomópontokat összekötő élekre természetesen felkerül a kapcsolat előjele és/vagy mértéke is (például az adott genetikai eltérés erősen csökkenti a beteg inzulinérzékenységét).



4. ábra • A Physiome-modell modu-láris felépítése

Bármely betegségmodell szélesebb körű klinikai alkalmazása előtt természetesen meg kell bizonyosodnunk arról, hogy az kellő pontossággal tükrözi a tényleges folyamatokat. Az Archimedes-modell segítségével például több tízezervirtuális cukorbetegben részletesen tanulmányozták a diéta, testmozgás és gyógyszeres terápia (például metformin) hatását, és azt találták, hogy a kapott szimulációs eredmények nagy pontossággal megközelítették a legkülönbözőbb klinikai vizsgálatokban kapott tényleges adatokat.

A top-down modellek másik előnye, hogy könnyebben figyelembe tudják venni az emberi szervezet komplexitását. A betegségek *in silico* modelljeiben ugyanis nem feleldehetünk meg a különböző szervek és szervrendszerek bonyolult együttműködéséről, és a homeosztatis és adaptációs reakciók sokaságáról sem. A betegségek komplex leírásához az élő szervezetben zajló többszintű szabályozási folyamatok megértésén keresztül vezet az út. A valósághű betegségmodelleknek tükrözniük kell a molekuláris, sejt- és szöveti szinten zajló anyag-, energia- és információcserét, továbbá a szervezet életfunkcióinak rendszerközi szabályozását végző neuroendokrin és immunrendszer működését is.

Elvileg a kétféle modellképzési eljárásnak valahol össze kell találnia, noha nemcsak kiindulási szintjük, de „filozófiájuk” is különbözik. Az alulról építkező modellek sokkal általánosabbak, a felülről indulók többnyire betegség-specifikusak; ezért egy bottom-up modellhez több top-down modell is csatlakozhat.

Döntéstámogató konzulensek

Egy probléma megoldásához gyakran az is elegendő, ha a klinikus gyorsan hozzá tud jutni a szükséges adatokhoz és/vagy kapcsolódó orvosi ismeretekhez. Ehhez különösen hasznosak az ún. ismeretszatók, amelyek egy adott problémához automatikusan ké-

pések előkeresni a megoldáshoz szükséges ismereteket anélkül, hogy a kereső feltételeket a klinikus előzetesen megfogalmazná. Sok esetben jól használhatók azok a kereső eszközök is, melyek a jelenlegihez hasonló szituációkat képesek visszakeresni különböző esettárakból. Az alternatív lehetőségek mérlegelésénél hasznos segítséget nyújthatnak a különböző számítógépes szimulátorok, melyek lehetővé teszik, hogy a klinikusok előzetesen megvizsgálják a szóba jövő beavatkozások várható hatásait.

A problémák megoldását közvetlenül is segítik a páciens-specifikus tanácsadásra alkalmas partnerrendszerek (Deutsch – Gergely, 2003). Ezek az elektronikus konzulensek egyaránt hozzájárulhatnak az állapot és kockázat felméréséhez, a diagnózis felállításához, és a legmegfelelőbb megelőzés és/vagy terápiás beavatkozás kiválasztásához. Ezek a konzulensek a jelenleg használt generikus betegségmenedzselő protokollokkal szemben az egyéni sajátosságokhoz illeszkedő tennivalókra vonatkozóan adnak útmutatást, és ezeket grafikus formában is megjelenítik a klinikusok számára.

Az egyes döntéstámogató szolgáltatásokat különböző módokon illeszthetjük be a genomikus medicina munkafolyamataiba. Az elektronikus konzulensek tanácsokkal segítik az orvosokat abban, hogy betegeiket a számukra legmegfelelőbb eljárásokkal gyógyítsák. A felügyelő programok folyamatosan figyelemmel kísérik az ellátás menetét, és üzenetekkel jelzik az elkövetett hibákat és mulasztásokat. Végül a kritikai programoktól azt várhatjuk el, hogy véleményt mondjanak az általunk helyesnek gondolt diagnózisról és/vagy terápiáról.

Következtetés

A legújabb orvosi és molekuláris biológiai kutatások eredményeként a jelenleg elterjedt betegségvezérelt, konfekcionált gyógyító módszereket fokozatosan az egyéni geneti-

kai eltérésekhez alkalmazkodó megelőző, szűrő- és gyógyító eljárások váltják fel. A kórfolyamatok molekuláris mechanizmusaira és az emberi genomon található mutációkra/polimorfizmusokra támaszkodó kockázatfelmérés és egészség-menedzselés feltételezi az orvosi informatikai és bioinformatikai módszerek integrált felhasználását. Az informatikai szolgáltatások sokfélék lehetnek.

Fontosak a különböző adatbázisok és esettárak, melyekben a klinikusok megtalálhatják a felmerülő problémák megoldásához szükséges adatokat és ismereteket. A tárolt adatok hatékony lekérdezéséhez standard ontológiákra, fogalomtárakra és keresőprogramokra van szükség, melyek a heterogén és elosztott információtárakból is képesek előkeresni a kért információt.

Az alternatív lehetőségek mérlegelésénél hasznosak lehetnek a számítógépes betegségszimulátorok, melyek segítségével kellő pontossággal előreláthatjuk, hogy az adott páciens miként fog reagálni a tervezett megelőző programokra és gyógyító beavatkozásokra, és ezáltal kiválaszthatjuk az optimális, egyénre szabott eljárást. A különböző döntéstámogató rendszerek közvetlenül és aktív módon vesznek részt a problémák megoldásában. A konzulens, felügyelő és kritikai programok egyaránt azt a célt szolgálják, hogy segítsék az állapot- és kockázatfelmérést, a diagnózis felállítását, továbbá a legmegfelelőbb megelőzés és terápiás beavatkozás kiválasztását. A rendszerek működését hatékony tudásmenedzselő eszközök biztosítják, melyek képesek feldolgozni a betegségek molekuláris mechanizmusaira vonatkozó különböző szintű ismereteket és az adott páciensre vonatkozó nagyszámú genomikai és klinikai adatot.

Az itt felvázolt korszerű orvoslás megvalósulásának több feltétele is van annak ellenére, hogy a közleményben leírt módszerek zöme már ma is létezik, vagy a viszonylag közeli jövőben valószínűleg létrejön. Az egyik alapvető feltétel a korszerű eszközpark létrehozása, ami vonatkozik a központi tároló és munkaállomásokra (például szerverrendszerekre, adattárakra, számítási és modellkészítő állomásokra) és a felhasználói eszközökre egyaránt. A probléma már ma sem az, hogy nincsenek ilyen eszközök, hanem az, hogy mikorra és hogyan terjednek majd el a genomikus medicina mindennapi alkalmazásához szükséges mértékben.

Hasonló korlátot jelentenek a felhasználók ismereteinek hiányosságai. Ezen a képzés révén is lehet segíteni, de sokat jelenthet az is, ha felhasználóbarát rendszereket fejlesztünk a gyakorló klinikusok számára. Leghatékonyabban a kettő kombinációja tűnik. Fontos hangsúlyozni, hogy a belátható jövőben még a legkorszerűbb informatikai rendszerek sem lesznek képesek teljesen kiszorítani az emberi tapasztalatot és intuíciót, azaz az orvos és egészségügyi személyzet részvételét a gyógyító folyamatban. Az emberi tényező szerepének hangsúlyozása mellett ugyanakkor ez a közlemény arra akarta felhívni a figyelmet, hogy a korszerű bioinformatikai és orvosi informatikai rendszerek hatékonyan segíthetik a gyógyító orvosok, ápolók és gondozók munkáját, és ezáltal hozzájárulhatnak az ellátás minőségének javulásához és a költséghatékony medicina széleskörű elterjedéséhez.

Kulcsszavak: *genomikus medicina, bioinformatika, orvosi informatika, kockázatmenedzselés, in silico modellek*

IRODALOM

- Bassingthwaighte, James B. (2000): Strategies for the Physiome Project. *Annals of Biomedical Engineering*. 28, 1043-1058
- Deutsch Tibor – Gergely Tamás (2003): *Kibermedicina. Medicina*, Budapest
- Guttmacher, Alan E. – Collins, Francis S. (2002): Genomic Medicine, a Primer. *New England Journal of Medicine*. 347, 19, 1512-1520
- Jeong, Hawoong – Tombor, B. – Albert, R. – Oltavi, Z. N. – Barabási, A. L. (2000): The Large-scale Organization of Metabolic Networks. *Nature*. 407, 6804, 651-654
- Kanehisa, Minoru – Goto, Susumu (2000): KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research*. 28, 27-30
- Martín-Sánchez, Fernando – Maojo, Victor – López Campos, Guillermo (2002): Integrating Genomics to Health Information Systems Methods. *Methods of Information in Medicine*. 41, 25-30
- Noble, Denis (2002): Modelling the Heart: From Genes to Cells to the Whole Organ. *Science*. 25, 1678-82
- Roses, Allen D. (2002): Pharmacogenetics Place in Modern Medical Science and Practice. *Life Sciences*. 15,1471-1480
- Sarachan, Brion D. – Simmons, Melvin K. et al. (2003): Combining Medical Informatics and Bioinformatics Toward Tools for Personalized Medicine. *Methods of Information in Medicine*. 12, 34-41
- Schlessinger, Leonard – Eddy David M. (2002): Archimedes: A New Model for Simulating the Health Care System: The Mathematical Formulation. *Journal of Biomedical Informatics*. 35, 1, 37-50
- Trimmer, Jeff (2001): Virtual Patients: Using Biosimulation to Understand Metabolism. American Diabetes Association 62nd Scientific Sessions. June 14-18. San Francisco
- Tomita, Masaru (2001): Whole Cell Simulation: A Grand Challenge of 21st Century. *Trends in Biotechnology*. 19, 6, 205-10

