

استانداردسازی داده‌ها در مطالعات بالینی داروها، نگاهی با ملاحظات CDISC

دکتر علی رزمی^۱، دکتر حسن ترکمندی^۲

۱. استادیار پژوهش - گروه فارماکولوژی و طب کاربردی مرکز تحقیقات گیاهان دارویی جهاد دانشگاهی - کرج - ایران
۲. داروساز عمومی - پژوهشگر بخش مراقبت‌های دارویی بیمارستان دکتر شریعتی - تهران

ایجاد کرد. یکی مطالعه افزایش خطر خودکشی در کودکان مصرف‌کننده داروهای ضدافسردگی و دیگری جمع‌آوری Vioxx از بازار دارویی. متعاقب این دو اتفاق، پزشکان، بیماران و حامیان شرکت‌های دارویی هر یک به نوبه خود آگاهی درباره خطرات نسخه‌نویسی، مصرف و فروش داروهای ضدافسردگی و ضدآرتروز را به یک‌باره حق قانونی خود دانستند! بار تشخیص این اثرات روی ناظران بالینی و آماری FDA بود، چراکه FDA تنها مرجعی بود که به تمام داده‌ها دسترسی داشت و باید تصمیمات آگاهانه‌ی در برابر مخاطبانش می‌گرفت. ثبت داده‌ها در شرایطی که در یک دسته دارویی بیش از ۱۰ ترکیب مختلف از چندین

دنیای فناوری اطلاعات یا همان IT با ورود در زندگی ما پیچیدگی‌های بیشتری را به زندگی امروزه اضافه کرده است. این پیچیدگی الزام تربیت افراد متخصص در هر رشته‌ای که IT در آن نفوذ کرده و همچنین تبیین اهمیت آن را برای جامعه هدف اجباری می‌سازد. امروزه داده‌های بالینی داروها به فراداده‌هایی (Metadata) تبدیل شده‌اند که تحلیل آن‌ها نیازمند کمک گرفتن از دنیای IT است. این مقاله سعی در بیان گوشه‌ای از نقش IT در ارایه و تحلیل داده‌های مطالعات بالینی داروها دارد.

دو رویداد مهم در زمینه ایمنی دارویی در سال ۲۰۰۴ تأثیر شگرفی بر روی ناظران FDA

۳ - مدل ODM (Operational Data Model)
یک مدل داده‌ها براساس XML² است که اجازه رد و بدل هر داده موجود در یک مطالعه بالینی را می‌دهد.

۴ - مدل Define.xml (Case Report Tabulation) (Data Definition Specification)
امروزه فایل‌های PDF دارند جای خود را به فایل‌های XML ارسال شده به FDA می‌دهند. Define.xml براساس ODM است و داده‌ها را برای کامپیوتر قابل خواندن می‌کند (برخلاف فایل‌های PDF). چون فایل XML برای کامپیوتر قابل خواندن است، تحلیل فراداده‌ها توسط نرم‌افزارها ممکن می‌شود.

با این مدل‌ها استخراج، وارد کردن و ایجاد داده‌ها ممکن می‌شود. پس یادگیری آن برای ارزیاب بالینی لازم است. FDA خواهان استفاده از مدل‌های داده‌ای در راهنمای خود است. با توجه به روند تغییرات ثبت داده‌ها در سال‌های گذشته پس از مدتی تنها شکل قابل ارایه به FDA مدل‌های CDISC خواهند بود (جدول ۱).

با این حال هنوز ثبت الکترونیک داده‌ها به صورت اختیاری انجام می‌گیرد. تلاش شد که بسیاری از اجزای آن در فاز ۱ مطالعات بالینی کنار گذاشته، تصمیم‌گیرندگان با عدم تطابق داده‌های اختصاصی هر مطالعه ارایه شده با راهنمای SDTM مشکل داشتند. بنابراین ایده ابتدایی CDISC کردن برای ثبت منظم، تراش زیادی خورده است. فاز مطالعه، بزرگی سازمان، چگونگی فهم استانداردها و میزان منابع عوامل این مطالعه تصمیم‌گیری هستند.

کارخانه تولیدکننده و اشکال دارویی مختلف کار آسانی نبود. تعجبی نداشت که کارشناسان FDA اذعان داشتند «اگر ما تمامی داده‌ها رو می‌خواهیم باید یک پایگاه داده عظیم داشته باشیم.» در اداره مواد غذایی و دارویی ایالات متحده آمریکا، تیم‌های مختلفی به منظور بررسی استانداردها و ایجاد یک زبان مشترک در ثبت داده‌های بالینی شکل گرفت و ماحصل آن به یک گروه غیرانتفاعی با نام اختصاری (Clinical Data Interchange Standards) CDISC Consortium ختم شد.

فعالیت اصلی CDISC به زبان ساده، استانداردسازی داده‌های بالینی کارخانه‌های داروسازی و ارایه مدل‌های داده‌ای مورد استفاده محققان این حوزه است. در ادامه به بررسی ۴ مدل رایج از مدل‌های داده‌ای CDISC پرداخته شده است:

۱ - مدل SDTM (Study Data Tabulation) (Model)

این مدل ارایه داده‌ها به صورت جداول مورد تایید FDA را تعریف می‌کند. FDA اعلام کرده است که SDTM را به صورت پرونده الکترونیکی^۱ [Electronic Common Technical (eCTD)] [Document case (CRTs)] قبول می‌کند. SDTM در اصل برای تسهیل جدول‌سازی مطالعات موردی [report tabulations] طراحی شده است و بنابراین SDTM به راحتی تهیه می‌شود.

۲ - مدل ADaM (Analysis Dataset Models)
مدل ADaM راهنمای تعریف داده‌ها را برای تحلیل آن داده‌ها تعریف می‌کند. این مجموعه داده برای خلاصه‌سازی به همراه تحلیل داده طراحی شده است.

جدول ۱ - تاریخچه جریان مستندات مورد نیاز FAD برای استانداردسازی داده‌ها

زمان	استانداردسازی انجام گرفته
جولای ۲۰۰۴	ارجاع مشخصات داده‌های مطالعه‌ای (study data specifications) eCTD برای جدول کردن داده‌ها به SDTM
مارس ۲۰۰۶	ایجاد استانداردهای داده‌ها
سپتامبر ۲۰۰۶	تکمیل بازنگری SDTM/ADaM و نمایش در CDISC
سپتامبر ۲۰۰۶	بسته شدن پرونده راهنمای (e-NDA) New Drug Application قدیمی ← eCTD دیگر تنها راهنمای ارایه داده‌ها شد.
دسامبر ۲۰۰۶	ارایه قوانین لازم برای ارایه داده‌های الکترونیکی در ثبت فدرال
مه ۲۰۰۸	درآمدن اولین ویراست PDUFA ³ IV IT ← افزایش بسیار اعتبار SDTM
اکتبر ۲۰۰۹	درآمدن ویراست ۱/۵ مشخصات داده‌های مطالعه‌ای، ایجاد مرجع مختص استانداردهای ADaM برای تحلیل داده‌ها
مارس ۲۰۱۰	درآمدن ویراست نخست نقشه استانداردسازی داده‌های CDER ← ایجاد اعتبار برای استانداردهای CDISC
مه ۲۰۱۱	درآمدن ویراست نخست (پرونده استانداردهای کلی داده‌های CDER) ← تشویق حامیان به ارایه استاندارد داده‌ها
فوریه ۲۰۱۲	درآمدن پیش‌نویس FDA برای «ثبت در قالب الکترونیک» (راهنمای e-Data) ← تشویق FDA برای ثبت الکترونیک داده‌ها

■ کدام مدل و کجا انتخاب شود؟

داشتن داده‌های جهانی استاندارد، کاربران را در تحلیل و بازنگری داده‌ها به پیش می‌برد. در یک مطالعه چه بازنگران پزشکی و آماری، چه اسپانسر و چه داور هماهنگی ناظر، از آشنا بودن به چگونگی سازماندهی داده‌های یک مطالعه سود می‌برند. این چنین مشخص می‌شود که داده‌ها مربوط به مطالعه غیربالیبی، داده‌های جدول شده SDTM یا فایل‌های تحلیلی ADaM هستند. کسی که مسوول استانداردسازی داده‌ها است

بیش از همه درگیر کار است. قبل از استانداردسازی داده‌ها، اختلاف بین دو نوع داده در داده خام و فایل آنالیز وجود داشت. بنابراین، تطابق ساختار و نام متغیرها با داده‌های اصلی نیازمند دستکاری جزئی داشت. امروز نیازی به نگرانی در مورد فرآیند طاققت‌فرسای تبدیل داده‌های خام به شکل SDTM و متعاقب آن ADaM نیست. برای سازمان‌های با حجم منابع زیاد، به کارگیری ADaM و SDTM مشکل نیست، بلکه سؤال مطرح برای آن‌ها این است که آیا در این فاز مطالعاتی

■ طراحی نقشه راه و وضع قوانین

برای انجام CDISC بهتر است نقشه دقیق راه از یک منبع داده به منبعی دیگر تعیین شود. این کار مزیت‌هایی دارد:

■ باعث می‌شود که داده‌ها از دست برنامه‌ریز به دست برنامه‌نویس برسد و دیگر مشکلات انجام هر دو توسط یک نفر وجود ندارد.

■ در اسرع وقت یک برنامه‌نویس روی داده‌ها بحث کرده و آن را مرور و تأیید می‌کند و تداخل کاری بین تصمیمات برنامه‌نویسان در یک پروژه حذف می‌شود.

■ محصول نهایی قابل بازبینی شده و در تمام قسمت‌های دیگر مسیر قابل ارجاع و تبدیل به داده‌های قبلی است.

کسی که روی CDISC کار کرده، می‌داند که عمده کار آن وارد کردن داده‌ها می‌باشد. تفاوت در شیوه‌ارایه داده‌ها از فردی به فرد دیگر و پافشاری آن‌ها بر شیوه خود بسیار اتفاق می‌افتد. به این منظور کار گروه‌ها در سازمان‌های زیادی قوانینی را تعیین کرده‌اند که مورد قبول تمام متخصصین این زمینه باشد. هم‌چنین آن‌ها برای مشابه‌سازی روند مرور و تأیید مستندات نیز به کار گرفته می‌شوند. برای سازمان‌های کوچک استفاده از کارگروه‌ها برای تدوین قوانین ممکن است غیرممکن باشد. آن‌ها می‌توانند از برنامه‌ریزی یک کارمند با سابقه یا مشاور خارجی سود ببرند. این‌گونه هر چند احتمال داشتن تداخل باقی است ولی حداقل می‌توانند از تصمیمات غلط ناشی از ساده کاری که منعکس در کل کار می‌شود، جلوگیری کنند.

این کار ارزش دارد؟ (هر چه به فاز اولیه نزدیکتر باشد احتمال پیشبرد تا ثبت دارو کمتر است) ولی برای بقیه روش نشستن و منتظر بودن با توجه به مهارت و منابع کم آن‌ها مناسب‌تر است. پس بهتر آن است که محصولاتی برای پیشبرد توسعه در این زمینه فراهم شوند.

■ آغاز با CDASH

بهترین راه برای شروع تطابق با سازمان CDISC وارد کردن اولیه استانداردهای داده‌های به دست مده است. Clinical Data Acquisition Standards Harmonization (CDASH) با عنوان FDA Consensus on Standards for Case Report Forms (ماده ۴۵) طراحی شده که بیشتر روی نام متغیرها تأکید دارد. هماهنگی ویژه‌شناسی با SDTM، راه را برای مرحله بعد آسان می‌کند. در ساختار CDASH دامنه‌های مورد هدف SDTM شامل عوارض دارو، اطلاعات فردی آزمایش شونده‌گان و داروهای دیگر که همزمان مصرف می‌شوند و غیره به صورت پیش‌فرض در آن قرار داده شده است. SDTM تعداد ۱۶ دامنه که در مطالعه بالینی دارو پر استفاده هستند را پوشش می‌دهد (برای نمونه رجوع شود به ویراست ۱/۱ <http://www.cdisc.org/cdash:SDADH>).

برای سازمان‌های قدیمی، بزرگ و یا حجم بالا که با سیستم‌های قدیمی مدیریت داده‌ها کار می‌کردند، انتقال سریع به CDASH ممکن است سودمند نباشد و این مسأله نیازمند کمی خلاقیت است که چگونه و چه زمانی شروع به CDISC کردن کنند.

■ ملاحظات SDTM

تصمیم‌گیری در مورد شروع SDTM آسان نیست. شروع SDTM در مرحله قبل از ثبت دارو (Pre-NDA) یعنی تبدیل داده‌های زیاد به کم و اگر صبر کنیم تا فاز III مطالعه که دیگر کار از کار گذشته است. از طرف دیگر، شروع در فاز I نیازمند کار زیادی برای دارویی که ممکن است اصلاً به فاز II و III نرسد. سازمانی که می‌خواهد SDTM را شروع کند باید به سؤالات زیر پاسخ دهد:

۱- آیا منابع و تبحر لازم برای SDTM را دارد؟ برای پاسخ به این سؤال باید به SDTM IG مراجعه کرده و جزییات و ساختار آن به دقت مورد مطالعه قرار داد.

۲- آیا مطالعه کافی برای یادگیری موثر مراحل نهایی انجام شده است که اجازه تبدیل داده به هر تحقیق را بدهد؟

مهارت با تمرین مستمر حاصل می‌شود. هر قدر که در کار وارد کردن داده‌ها وقفه کمتری روی دهد شانس فراموشی کارهای انجام شده کمتر است.

۳- آیا فضای ثابت برای اتوماسیون تبدیل بخش‌های مشخص از یک مطالعه به مطالعه دیگر را دارید؟

اولین تبدیل داده‌ها یک تجربه یادگیری است. در مطالعات بعدی تشابه داده‌های خام با داده‌های تجربه شده قبلی فرصت‌هایی مثل استفاده از نرم‌افزارهای بیشتر را برای بهتر شدن کار شما فراهم می‌کند.

۴- آیا برنامه‌ریزی برای استفاده از همه ابزارهایی که در مراحل پاکسازی داده‌ها و آنالیز درست فرآیند SDTM به کار می‌آید، انجام داده‌اید؟ به عنوان مثال اگر می‌خواهید از پروفایل بیمار

گزارش تهیه کنید SDTM برای این کار ابزار مفیدی دارد؟

۵- در چه فازی از مطالعه بالینی هستید؟ راهنماهای زیادی برای چگونگی مشارکت دادن داده‌ها درست شده‌اند. داده‌های مطالعات بالینی فاز I که با مقادیر مصرف متفاوت روی افراد سالم چه از منظر فارماکوکینتیک و چه فارماکودینامیکی انجام شده است، قابل ادغام با داده‌های فاز II و III نیستند. داده‌های مطالعات فازهای نهایی ارزش بیشتری در ارائه SDTM دارند.

۶- آیا باید یک نگرش وابسته به مرحله فاز داشت؟

ممکن است شما یا سازمان شما منابع یا مهارت لازم برای تبدیل کامل SDTM را نداشته باشید. با این حال برخی از دامنه‌های مورد پوشش آن مثل عوارض جانبی، اطلاعات شخصی، داروهایی که همزمان مصرف می‌شود یا اطلاعات آزمایشگاهی می‌تواند به درد بخورد. این‌گونه اجرای یک برنامه استاندارد تسهیل می‌شود و شما می‌توانید از ابزارهای معینی برای تبدیل در آینده استفاده کنید. هر چند که به درد ثبت در FDA نمی‌خورد.

■ ملاحظات ADaM

ADaM برای دو هدف درست شد:

تعریف یک ساختار استاندارد برای تحلیل‌های معمول مورد نیاز

به وجود آوردن راهنمای ارائه ADaM

آخرین ویراست ADaM در سال ۲۰۰۹ تدوین شد، برای دیدن نمونه رجوع شود به:

<http://www.cdisc.org/adam>

در سناریو دوم احتمال عدم همخوانی داده‌ها با تحلیل‌های قبلی ADaM وجود دارد. بعضی مواقع ساختن دوباره تحلیل‌های ADaM به صرفه است که نیازمند دوباره کاری است. هم‌چنین ADaM ایجاد شده ممکن است با مدل جدید آن تطابق نداشته باشد که نیازمند بازبینی دوباره داده‌های بزرگ و SDTM است.

در این‌جا لازم است از گروه‌های دیگری که علاوه بر CDISC برای استانداردسازی داده‌های بالینی ایجاد شده‌اند، نیز نام برده شوند:

- International Conference on Harmonization, at <http://www.ifpam.org>
- Society for Clinical Data Management, at <http://www.scdm.org>
- Association for Clinical Data Management, at <http://www.acdm.org.uk>
- Drug Information Association, at <http://www.diahome.org>
- Applied Clinical Trials, at <http://www.actmagazine.com>
- Statisticians in the Pharmaceutical Industry (PSI), at <http://www.psiweb.org>

شامل داده‌ها به همراه تحلیل‌های آماری است که به ناظران پزشکی و آماری اجازه تأیید و ارزیابی سلامت داده‌ها را می‌دهد. این چنین ارتباطی بین داده‌های اصلی و تحلیل‌ها برقرار می‌شود. در ابتدا باید مشخص شود که منبع داده‌ها کجا است و چه قسمتی از آن قرار است، تحلیل شود. باید توازن علیه داده‌های کم فارماکوکینتیک و مطالعات ایمنی مربوط به فاز ۱ که هستند، اعمال شود. به دلیل این‌که: ۱- تحلیل‌های روی این داده‌ها بسیار پایه‌ای هستند و ۲- معمولاً ناظرین توقع ثبت آنالیز داده‌های آن‌ها ندارند. ایجاد یک دسته از داده‌های ADSL که شامل یک داده ثبتی برای هر فرد مورد آزمایش است، حداقل نیاز داده‌ای مورد نیاز برای ADaM است.

در اکثر پژوهش‌سکده‌ها برای انجام CDISC نیازمندی به خلاصه‌سازی داده‌های خام گزارش‌های بالینی به منظور آماده شدن برای تحلیل در مرحله بعدی است. یک سناریو عمل کردن به صورت سری است و دیگری انجام هر دو همزمان با هم؛ یعنی دستیابی به نتایج تحلیل در زمان کمتر.

منابع

1. Holland C. Shostak J. Implementing CDISC Using SAS®; An End-to-End Guide: SAS Institute Incorporation; 2012: 7-19.
2. SAS Institute Inc. SAS® 9.2 CDISC Procedure: User's Guide. Cary, NC: SAS Institute Inc; 2008: 3.
3. Shostak J. SAS® Programming in the Pharmaceutical Industry: SAS Institute Incorporation; 2005: 2-6.

زیرنویس‌ها

۱. eCTD ساختار الکترونیکی آرایه داده‌ها در FDA است.
۲. Extensible Markup Language (XML) نوعی کدنویسی است که هم برای انسان و هم برای کامپیوتر قابل خواندن است.
3. PDUFA: Prescription Drug User Fee Act (puh-DOO-fa)