

چگونه می‌توان مقالات مروری پزشکی درباره مؤثر بودن درمان را مطالعه کرد؟

عباس عرب ورامینی* مهییز رهبر آزاد**

چکیده

بسیاری از تصمیم‌گیرها در مورد ارائه بهترین درمان و تستهای تشخیصی در بیمارستانها و کلینیکها بر اساس شواهد علمی روز گرفته نمی‌شود. در عین حال، منابع موجود جهت تصمیم‌گیریهای کلینیکی قابل اعتماد نبوده و یا در بسیاری از موارد توسط متخصصین مختلف مورد تناقض می‌باشد. در چنین محیطی مسئولیت قضاوت کلینیکی متخصصین بر عهده خود آنان می‌باشد. وجود انحراف (*Bias*) در بسیاری از تحقیقات دارا بودن تواناییهای فراتر از جستجوی تنها در بانکهای اطلاعاتی را برای درمانگرها اجتناب‌ناپذیر کرده است. درمانگرها باید مجهز به دانش تجزیه و تحلیل تحقیقات علمی شوند. این دانش توسط پارادایمی جدید در علوم پزشکی تحت عنوان *Evidence - Based Practice* امکان‌پذیر گشته است. حجم وسیع مقالات چاپ شده در زمینه علوم پزشکی دسترسی و قضاوت در مورد بهترین اطلاعات را غیرممکن ساخته است. بکارگیری *Check list* می‌تواند فرد را در جهت انجام سریع و سیستماتیک و تجزیه و تحلیل اطلاعات کمک کند. هدف این مقاله نشان دادن راهکارهایی سیستماتیک جهت ارزیابی تحقیقات علمی در زمینه مؤثر بودن درمان (*Treatment efficiency*) می‌باشد (مجله طبیب شرق، سال سوم، شماره ۲، ص ۱۱۱ تا ۱۲۲، تابستان ۱۳۸۰).

مقدمه

تولد پارادایمی جدید به نام *Evidence - Based Medicine (EBM)*

Case I: مرد ۵۰ ساله‌ای در یکی از شهرستانها در آخر هفته به دلیل گیر افتادن انگشتش زیر لولای چرخ و صدمه به بند انتهائی به کلینیک شما مراجعه می‌کند. در معاینه معلوم می‌گردد که انگشت دچار *Mallet-Finger* شده است و قادر به *Extension* بند انتهائی آن نمی‌باشد. عکس رادیوگرافی هیچ گونه شکستگی را تایید نمی‌کند. شما آگاه هستید که در این مورد دو نظر متناقض، شامل درمان محافظه کارانه با گرفتن اسپلینت و جراحی وجود دارد. شما با جراح ارتوپد *On call* تماس گرفته و شرایط فوق را توضیح می‌دهید. او نظر می‌دهد که این مشکل معمولاً با اسپلینت خوب می‌گردد و موارد کمی احتیاج به جراحی

دارند. بیمار از توضیح فوق راضی نمی‌شود و درخواست یک نظر دوم میکند. در تماس دوم با جراح ارتوپد دیگر او نظر می‌دهد که تقریباً هیچ کدام از بیماران من با اسپلینت نتیجه خوبی نشان نداده و احتیاج به جراحی داشته‌اند. *Case II*: فرض کنید یک مرد ۲۳ ساله بدنبال پیچ خوردگی مچ پا در فوتبال به شما مراجعه می‌کند. شما به طور معمول درخواست *X-Ray* می‌کنید. عکس رادیوگرافی هیچ گونه شکستگی نشان نمی‌دهد. بر همین اساس شما درمان معمول شامل استراحت، یخ درمانی، بانداژ فشاری و بالا نگهداشتن اندام را توصیه می‌کنید. شاید در اولین نگاه این کار صحیح‌ترین و بی خطرترین راه برخورد با چنین بیماری بنظر رسد و شما هر روز آن را بارها انجام دهید (*Subjective Judgement*) اما با کمی تأمل ممکن است

* دانشگاه علوم پزشکی و خدمات بهداشتی درمانی زاهدان، دانشکده پزشکی، گروه فیزیوتراپی
** دانشگاه علوم پزشکی و خدمات بهداشتی درمانی شهید بهشتی، بیمارستان مسیح دانشوری، دفتر پرستاری

است (*Peer review*) یا توسط شرکتهای تجاری ذینفع مورد حمایت قرار گرفته است. حتی مقالاتی که در آبرومندترین مجلات پزشکی چاپ می‌شوند کامل نمی‌باشند و دارای نقص در طراحی و تجزیه و تحلیل هستند.^(۲) موارد فوق سرعت پیشرفت مقالات پزشکی را کاهش داده و روند کیفی آن را در طی سه دهه اخیر در کمتر از ۵۰ درصد نگه داشته است.^(۳) فرایند فوق باعث به کارگیری تستها و روشهای درمانی گوناگونی گشته است که هنوز مؤثر یا مضر بودن آن توسط شواهد علمی معتبر تأیید نشده است.

به همین دلیل یک خواننده مقاله باید مسئولیت شخصی قضاوت در اعتبار و اهمیت کلینیکی یک مقاله را به عهده بگیرد. این بدین معنی است که کلیه درمانگرها و کسانی که کار کلینیکی انجام می‌دهند از یک طرف باید از آخرین یافته‌های علمی آگاه باشند و از طرف دیگر کیفیت آنها را بتوانند مورد نقد و قضاوت قرار دهند. موارد فوق ضرورت ظهور یک سیستم جهت قبول یا رد یافته‌های علمی را بوجود آورده است. شکل‌گیری این سیستم به اواسط قرن ۱۹ بر میگردد، که در دو دهه اخیر بصورت مدون و تحت عنوان *Evidence - Based Medicine (EBM)* مورد توجه قرار گرفته است.

EBM: یک روش در یک سیستم بهداشتی است که جمع آوری، تفسیر و بهم مربوط ساختن اطلاعات معتبر بدست آمده از آخرین دستاوردهای تحقیقاتی را امکانپذیر می‌سازد.^(۱) در سیستم فوق ترکیبی از شواهد تحقیقاتی، تخصص کلینیکی و خواسته‌های بیمار مورد استفاده قرار می‌گیرد^(۴) تا شواهد علمی موجود را به شکل فرمول‌های قابل استفاده در کارهای روزمره کلینیکی درآورد.

کاهش شکاف بین یافته‌های تحقیقاتی و کاربری کلینیکی آن در کشورهای جهان

نگاهی به تاریخچه شواهد علمی در تصمیمات کلینیکی نشان می‌دهد که علیرغم نظرات عده‌ای از صاحب

شما بخواهید بدانید که بهترین راه تشخیص اینکه یک بیمار باپیچ خوردگی مچ چه وقت احتیاج به عکس رادیوگرافی دارد چیست؟ شما در *MEDLINE* جستجو میکنید و به تحقیقی برخورد می‌کنید که بطور همزمان در چندین مرکز انجام شده است و به عنوان *Ottawa ankle rules* شناخته می‌شود.^(۳) نتایج فوق نشان می‌دهد که تکیه بر قوانین فوق نتوانسته بدون افزایش تعداد شکستگی کشف نشده (*Missed fracture*) از تعداد عکس رادیوگرافی به مقدار قابل ملاحظه‌ای بکاهد. نتایج این تحقیق منجر به کاهش هزینه و صرفه جویی در وقت بیمار گشته است. شما آن را در مرکز خود بکار می‌بندید.

درمانگرها تمایل دارند که تصمیمات کلینیکی و درمانی خود را بر پایه موارد مختلفی چون تجربیات گذشته، درمان رایج در سایر مراکز، نظر سایر متخصصان و آموزشهای حرفه‌ای خود بگیرند. متأسفانه در حال حاضر این منابع کمتر قابل اعتماد می‌باشند چراکه موارد متعددی از تناقض در نظرات دو متخصص فعال در یک رشته خاص مشاهده می‌شود (بیمار I)، ثانیاً در بسیاری از این تصمیمات قضاوت نظری درمانگر جایگزین یک سیستم تجزیه و تحلیل متکی به شواهد علمی شده است (بیمار II)، از طرف دیگر تکیه بر تجزیه و اطلاعات سنتی و حتی کتابهای مرجع (*Text book*) نیز نمی‌تواند گره گشا باشد.^(۲) شاید بتوان این را به رشد خارج از تصور اطلاعات چاپ بیش از ۶ میلیون مقاله پزشکی در سال نسبت داد. کتابهای مرجع حتی در زمان انتشار یا مدت کوتاهی پس از آن به صورت تاریخ گذشته در می‌آیند.^(۱) این حجم از اطلاعات درمانگرها را روبروی جنگلی از اطلاعات در زمینه تصمیمات کلینیکی و درمانی قرار داده است. اگر درمانگرها سعی کنند با این رشد همراه شوند مثلاً "روزی حداقل دو مقاله بخوانند، در طی یکسال بیش از ۸۲ قرن از مطالب روز عقب می‌افتند. با این حال، علی‌رغم حجم غیر قابل کنترل مقالات پزشکی کمتر از ۱۵ درصد از این مقالات مفید می‌باشند.^(۱۲) تعداد زیادی از این مقالات توسط یک متخصص در همان رشته نقد و بررسی نشده

منظور بررسی گزارشات متناقض در ۱۹۹۵ یک تیم پزشکی از اکسفورد تمام درمانهای ارائه شده به ۱۰۹ بیمار را طی یکماه مورد بررسی قرار داد. (۷) سپس با جستجو در بانکهای اطلاعاتی، موثر بودن درمان را بر اساس شواهد علمی بررسی کردند. نتایج تحقیق فوق نشان داد که ۵۳ درصد از درمانها بر اساس RCT (شواهد درجه I) و ۲۹ درصد بر اساس تحقیقات غیر تجربی (شواهد درجه II) انجام پذیرفته است. (۷)

نظران (۴)، شکاف موجود بین این دو رویه کاهش است. گزارشات موجود نشان می‌دهد که در دهه ۷۰ تا اوائل ۸۰ فقط ۱۰ تا ۲۰ درصد از روشهای درمانی رایج در مراکز درمانی توسط شواهد علمی درجه I مانند کارآزمایی کنترل شده تصادفی (RCT) مورد تأیید بوده است. (۴) به نظر می‌رسد که میزان فوق در اوائل دهه ۹۰ بهبود یافته بطوری که حدود ۵۰ درصد از تصمیمات درمانی بر اساس شواهد علمی درجه I و II گرفته شده است (جدول ۱). (۶) به

جدول ۱- بکارگیری شواهد تحقیقی در کلیه خدمات بهداشتی درمانی آمریکا سال ۱۹۹۲

1-Strong evidence	+	Strong	consensus	2%
2-	"	Modest	"	2%
3-	"	Weak	"	5%
4-Modest evidence	+	Strong	consensus	20%
5-	"	Modest	"	25%
6- No/Weak	+	Strong	"	20%
7-	"	Modest	"	25%
8-	"	No/Weak	"	6%

علمی در سالهای اخیر غیر قابل انکار می‌باشد.

چگونه می‌توان یک مقاله را از نظر مؤثر بودن درمان ارزیابی کرد؟

تصمیم در مورد موثر بودن نتایج یک تحقیق و مقایسه آن با سایر روشهای درمانی متداول نیاز به دانشی فراتر از قدرت جستجو در بانکهای اطلاعاتی (Data bases) و مهارت در تجزیه و تحلیل اطلاعات دارد. فرایند ارزیابی یک مقاله در سه مرحله انجام می‌گیرد (۲):

۱- تحقیق در مورد اعتبار مقاله

۲- تعیین هدف مقاله

این در حالی است که فقط ۱۸ درصد درمانها توسط تحقیقات علمی تأیید نگردید. یک سال بعد تیمی از ژاپن گزارش داد که ۸۱ درصد از توصیه‌های داروئی بر اساس شواهد علمی درجه I و II بوده است (RCT: ۲۹٪). تحقیقات غیر تجربی: ۶۰٪). اخیراً نیز در تحقیق مشابه دیگری، ۸۴ درصد درمانهای ارائه شده در بیمارستانها یا توسط RCT (۵۰٪) تأیید گردیده است یا مورد موافقت بیشتر متخصصین بوده است (۳۴٪) و فقط ۱۲ درصد درمانها توسط این دو تأیید نگردیده است. (۹) علی‌رغم اینکه همه گزارشها به خوشبینی نتایج فوق نمی‌باشد، اما نیاز متخصصان و درمانگرها به کسب آخرین یافته‌های

۳- بررسی اعتبار مقاله بر اساس هدف آن

مرحله اول: اعتبار اولیه مقاله:

در قدم اول وقتی که به یک مقاله نگاه می‌کنید باید از خودتان بپرسید آیا این مقاله ارزش وقت گذاشتن به منظور مطالعه عمیق را دارد؟ این سوال در چند ثانیه قابل جواب دادن می‌باشد و فقط لازم است ۶ سوال ساده را از خود بپرسید. هرگونه توقف یا تامل در پاسخ به هرکدام از این سوالات می‌تواند شما را از گذاشتن وقت روی تحقیق مورد نظر منصرف کند. (۲)

سوال اول: آیا این مجله از آن دسته است که مقاله‌ها توسط متخصصان هم رشته (Peer-review) مورد ارزیابی قرار می‌گیرد؟

بسیاری از مجلات تخصصی باید از این مشخصه Peer-review برخوردار باشند. اگر نسبت به این موضوع شک دارید پاسخ این سؤال در داخل مجله در قسمت دستورات مربوط به فرستنده مقالات موجود می‌باشد. بطور معمول در مجلات معتبر وقتی مقاله‌ای دریافت می‌شود، ابتدا سردبیر تعیین می‌کند که آیا مقاله مناسب مجله مورد نظر می‌باشد و سپس در صورت مثبت بودن آن را برای چندین صاحب نظر جهت تجزیه و تحلیل می‌فرستد. این افراد در زمینه مورد نظر و روشهای تحقیقی بکارگرفته نظر می‌دهند. در یک بررسی گزارش شده که از ۱۲۰۰ مقاله‌ای که در سال به مجله *Annals of Internal medicine* می‌رسد نیمی توسط سردبیر پس فرستاده می‌شود و فقط ۱۵ درصد پس از بررسی توسط دو متخصص، چاپ می‌گردد. (۱۴)

علی رغم اینکه بکارگیری این روش کیفیت مقاله را بسیار بالا می‌برد اما تضمین نمی‌کند که مقاله چاپ شده بدون هرگونه اشتباه یا انحرافی باشد.

سوال ۲: آیا امکان دارد نمونه‌های انتخاب شده برای تحقیق از نظر جغرافیایی به نوعی با بیماران ما در منطقه‌ای دیگر از نظر واکنش به درمان متفاوت باشند؟ معمولاً در انتهای صفحه اول مقاله، مکان انجام تحقیق ذکر شده است. با اینکه ممکن است هیچ دلیلی جهت

متفاوت بودن بیماران شما با نمونه‌های موجود در مقاله وجود نداشته باشد، اما این سوال در ذهن شما شکل می‌گیرد که آیا این نتایج قابل تعمیم به بیماران شما می‌باشد؟

سوال ۳: آیا هزینه تحقیق توسط شرکت ساخت وسایل پزشکی یا شرکت دارویی که به نوعی در نتایج ذینفع می‌باشد تأمین شده است؟

صرف وجود چنین مسأله‌ای دال بر وجود انحراف (Bias) نمی‌باشد بلکه برعکس بسیاری از تحقیقات با طراحی خوب توسط شرکتهای دارویی یا پزشکی انجام گرفته است. نکته‌ای که در اینجا مطرح می‌باشد آوردن دلایل و نشانه‌هایی مطمئن از طرف نویسنده است که احتمال تأثیر پذیری نتایج از این شرکتها را رد کند.

سه سوال بعد به ارتباط کلینیکی تحقیق با حیطه کاری مورد علاقه شما می‌پردازد. بدست آوردن اطلاعات فوق با خواندن نتیجه مقاله و قسمت‌های انتخاب شده‌ای از چکیده مقاله حاصل می‌شود. با توجه به حجم انبوه مقالات منتشر شده، خواندن مقاله‌ای در ارتباط با یک بیماری نادر یا یک روش پیشرفته آزمایشگاهی که هرگز با آن روبرو نمی‌شویم یا در دسترس ما نمی‌باشد، تلف کردن وقت می‌باشد.

سوال ۴: آیا نتایج بدست آمده (اگر واقعی باشد) می‌تواند تأثیر مستقیم روی سلامت بیماران من بگذارد؟

شما جمله زیر را در نتیجه‌گیری یک مقاله می‌خوانید: استفاده از دکزامتازون خوراکی به خاطر کاربرد آسانتر آن، قیمت ارزاتر و قابل دسترس بودن آن نسبت به سایر روشها، ترجیح داده می‌شود. شما به چکیده مقاله توجه می‌کنید و متوجه می‌شوید که نتایج فوق بر اساس عواملی چون کاهش مدت زمان حضور در بیمارستان و مدت زمان سپری در اورژانس، تغییر علائم موجود در طی یک هفته و دفعات مراجعه به بیمارستان گرفته شده است. همه اینها از جمله مواردی است که شما در مداوای بیمارانتان به آن اهمیت می‌دهید، پس به این سوال جواب مثبت دهید.

سوال ۵: آیا مشکلی که توسط این مقاله پوشش داده شده است در حیطه کاری من شایع است و آیا وسیله یا روشی

حیطه کاری شما نیستند و یا کمتر روی روش بیمار دیدن شما تاثیر دارد.^(۲)

مرحله دوم: تعیین هدف یک مقاله:

بعد از پاسخ دادن به همه سوالات مرحله اول، وظیفه بعدی شما تعیین ضرورت انجام این تحقیق و نوع مشکل کلینیکی است که محقق در این مطالعه به آن جواب مثبت داده است.^(۱۵)

مقالات معمولاً به یکی از چهار طبقه زیر مربوط می‌باشند: ۱- درمانی ۲- تست‌های تشخیصی ۳- علتی ۴- پیش‌آگهی (Prognosis) که هر کدام ترجیحاً نوع خاصی از روش تحقیق را نیاز دارند (جدول ۲).

که فرد بکار برده است احتمالاً در دسترس من می‌باشد؟ اگر به عنوان مثال در حیطه کاری خود کمتر با نوزادان سروکار دارید و مقاله در این ارتباط باشد، شما به این سوال جواب منفی داده و به سراغ مقاله‌ای دیگر می‌روید.

سوال ۶: آیا نتایج فوق (فرضا حقیقی باشد) منجر به تغییر در روش درمانی من می‌گردد؟

در کمتر از چند ثانیه شما می‌توانید به این شش سوال پاسخ دهید. پاسخ مثبت به همه سوالهای فوق به شما اجازه می‌دهد که تصمیم بگیرید که آیا می‌خواهید به منظور بررسی تحلیل گرانه این مقاله وقت صرف نمایید یا نه؟ این روش امکان صرف نظر کردن از مقالاتی را می‌دهد که در

جدول ۲: روشهای تحقیقی متناسب با هدف مقاله (۱۵)

روش تحقیق	طبقه‌بندی
<i>RCT, Double blind, Placebo- Controlled</i>	تست مؤثر بودن یک درمان. دارو یا روش <i>Therapy:</i>
<i>Cross sectional survey</i> . مقایسه تست جدید با استاندارد	بررسی اعتبار و تکرارپذیری تست‌های تشخیصی <i>Diagnosis:</i>
<i>Cohort or case-control</i>	بررسی ارتباط یک فاکتور با شیوع یک بیماری <i>Causation:</i>
<i>Longitudinal cohort study</i>	تعیین نتایج یک بیماری <i>Prognosis:</i>

جهت شروع این مرحله، ابتدا تمام چکیده را بخوانید سپس نگاهی به عنوان‌ها، جداول، شکلها و جملات اول پاراگراف مقاله بیندازید تا ساختار کلی مقاله را بدست آورید. بدنبال این کار تمام توجه شما باید معطوف به قسمت روش تحقیق مقاله گردد. جهت یک مقاله در ارتباط با موثر بودن درمان شما باید به سوالهای اختصاصی درج در جدول ۳ پاسخ گوید.

آیا تحقیق *RCT* بوده است؟ کارآزمایی تصادفی کنترل شده (*RCT*) در روشهای تحقیقی به عنوان استاندارد طلایی جهت تعیین موثر بودن درمان در نظر گرفته می‌شود.^(۲) قدرت *RCT* در استفاده از روش انتخاب تصادفی نمونه نهفته است. به طور کلی آزمونهای غیر تصادفی تمایل به

پاسخ به مرحله شماره ۲ با خواندن چکیده مقاله امکان پذیر می‌باشد. همینطور مراجعه به قسمت معرفی مقاله به خصوص مراجعه به پاراگراف آخر، هدف یک مقاله را تعیین می‌کند و در صورتی که با حیطه کاری شما مطابقت داشت شروع به خواندن و تجزیه و تحلیل آن مقاله کنید.

مرحله سوم: بررسی اعتبار مقاله بر اساس هدف آن:

بعد از پشت سر گذاشتن موفقیت آمیز دو مرحله قبلی، مقاله باید به صورت انتقادی و در ارتباط با اعتبار و کاربری آن در حیطه کار شما مورد تجزیه و تحلیل قرار گیرد. همانطور که در بالا توضیح داده شد هر یک از چهار طبقه بندی که در جدول آمده است ترجیحاً دارای یک روش تحقیقی مخصوص به خود می‌باشد که اعتبار آن را تضمین می‌کند.

جدول ۳- چک لیست مربوط به اعتبار یک مقاله در ارتباط با درمان

- ۱- آیا تحقیق RCT بوده است؟ بله خیر چطور نمونه‌ها انتخاب شده‌اند؟ با توضیح.
- ۲- آیا نمونه‌های شرکت کننده در تحقیق مشابه بیمارانی است که خواننده با آن روبرو می‌باشد؟ بله خیر
- ۳- آیا کلیه کسانی که در تحقیق شرکت کرده‌اند در تجزیه و تحلیل و نتیجه‌گیری شرکت داده شده‌اند؟ آیا Follow-up کامل بوده است و تعداد کمی در انتهای کار حضور نداشته‌اند؟ بله خیر
- ۴- آیا کلیه کسانی که در تحقیق شرکت داشته‌اند (نمونه و محقق) نسبت به درمان بی اطلاع بوده‌اند؟ بله خیر
- ۵- آیا گروه کنترل و شاهد در شروع تحقیق مشابه بوده‌اند؟ بله خیر
- ۶- آیا دو گروه به صورت مساوی مورد درمان قرار گرفته‌اند؟ بله خیر
- ۷- آیا نتایج همانطور که از نظر آماری معنی دار هستند از نظر کلینیکی نیز مهم می‌باشند؟ بله خیر
- ۸- اگر نتیجه یک کارآزمایی منفی است آیا آزمون قدرت (Power) انجام گرفته است؟ بله خیر
- ۹- آیا سایر عواملی که می‌توانند در نتیجه‌گیری تأثیر گذارند، گزارش شده‌اند؟ بله خیر
- ۱۰- آیا فواید درمان به احتمال خطرات آن برتری دارد و مقرون به صرفه است؟ و آیا می‌توان آن را به بیماران خود تعمیم داد؟ بله خیر

اشاره کرده باشد. تکنیکهای مختلفی جهت تصادفی سازی مورد استفاده قرار می‌گیرد. محقق ممکن است از روش تصادفی سازی استفاده کرده باشد بدین شکل که هر شرکت کننده دارای شانس مساوی جهت ثبت نام در یک گروه یا گروه دیگر باشد بدون توجه به اینکه نفرات قبلی در چه گروهی قرار گرفته‌اند. این نوع روش ممکن است باعث افزایش تعداد نفرات یک گروه نسبت به گروه دیگری شود یا منجر به اختلافاتی اصولی در دو گروه شود که بر نتایج درمان تأثیر می‌گذارد. جهت اجتناب از مشکلات فوق می‌توان از blocked Randomization استفاده کرد که تعداد نمونه‌ها در هر گروه مساوی است یا Stratified Randomization که در آن نمونه‌ها در داخل گروه‌ها بر اساس عوامل مشکل ساز نهفته مانند سن یا جنس به صورت تصادفی قرار می‌گیرند.^(۷)

در ارزشیابی یک تحقیق که نمونه‌ها به صورت تصادفی انتخاب شده‌اند مثلاً از روش Case-control یا Cohort analysis استفاده شده باشد، می‌توان از قانون اندازه‌گیری اطمینان که توسط Sackett ابداع شده استفاده کرد تا میزان

نشان دادن نتایج دلخواه‌تری نسبت به مطالعات تصادفی دارند و این منجر به ایجاد انحراف (bias) در جهت درمان جدید می‌شود به خصوص اگر گروه کنترل یا گروه درمان با Placebo وجود نداشته باشد.^(۵)

بر همین اساس شاید بتوان گفت که حقیقت بهتر بودن تصادفی سازی در یک درمان جدید ناشی از این فرض می‌باشد که مشخصات شناخته شده و ناشناخته بیماران مورد بررسی با شانس مساوی بین دو گروه درمان و کنترل توزیع شده‌اند. در واقع نتایج کلینیکی بدست آمده در حین تحقیق می‌تواند ناشی از عوامل گوناگونی باشد که خود درمان یکی از آنهاست. عواملی چون سیر طبیعی بیماری، شدت، بیماری، بیماریهای همراه، سن، جنس، وضعیت اجتماعی و روانی و غیره می‌تواند به صورت مستقیم و یا غیر مستقیم روی نتایج درمان تأثیر گذار باشند.^(۵)

به همین خاطر تحقیقات غیر تصادفی در تعیین موثر بودن درمان با محدودیتهای متعددی روبرو می‌باشند. یک مقاله در ارتباط با درمان در قسمت روش تحقیق باید به انتخاب تصادفی بیماران در هر گروه و نحوه تصادفی سازی آنها

به درجه‌های A, B, C تقسیم بندی کرد که در جدول ۵ آورده شده است.

قدرت یک روش تحقیقی را بدست آورد (جدول ۴). (۱۷)
همچنین بر اساس روشی دیگر، Sackett شواهد علمی را

جدول ۴: ارزشیابی اطمینان تحقیقات بر اساس روش تحقیق مورد استفاده (۱۷)

Level	Evidence	Certainty Scale
I	Large RCT with low false-positive and low false-negative errors	100%
II	Small RCT with high false-positive and high false-negative	
III	Non-randomization, Cohort analytical	
IV	Non-Randomization, Historical Cohort, Case control study	
V	Before - After Study, Descriptive study case series without control	0%

جدول ۵: قدرت شواهد علمی (۱۷)

توصیه به استفاده از شواهدی که نتایج درمانی آن بوسیله حداقل یک مورد تحقیق با <i>Level I</i> (یا بیشتر از یک مورد) تأیید شده باشد.	Grade A
توصیه به استفاده از شواهدی که نتایج درمانی آن بوسیله روش تحقیق با <i>Level II</i> تأیید شده باشد.	Grade B
شواهدی که توسط یکی از روشهای تحقیق <i>V, IV, III</i> انجام گرفته باشد و دارای ارزش علمی پایین تری می‌باشند.	Grade C

به بیماران کمردردی مراجعه کننده به شمادر ایران تعمیم داد. در این حال باید به امکانات رفاهی در محل کار، سطح آگاهی و آموزش، تفاوت نحوه کار، عوامل اقتصادی و روحی، سن، سابقه ورزشی و... توجه داشت. مثال دیگر اینکه شما در محل کار خود با بیماران بچه ۳ ماه تا ۵ سال مبتلا به سرخک روبرو هستید و تحقیق مورد مطالعه شما بر روی همین گروه سنی بوده است. شما بعد از توجه به شرایط حذف نمونه‌ها (*Exclusion criteria*) می‌توانید از

سؤال ۲: آیا نمونه‌های شرکت کننده در تحقیق، مشابه بیمارانی است که خواننده با آن روبرو می‌باشد؟
جهت تعمیم دادن نتایج یک تحقیق به آنچه ما در محل کار خود انجام می‌دهیم باید به شرایط انتخاب نمونه‌ها (*Inclusion criteria*) و شرایط حذف آنها از تحقیق (*Exclusion criteria*) در یک مقاله در ارتباط با درمان توجه داشته باشیم. مثلاً آیا می‌توان نتایج یک تحقیق که بر روی علل کمردرد شغلی در مردان امریکا پرداخته است را

نتایج تحقیق فوق استفاده کنید.

گاهی ممکن است بیماریارانی که در یک تحقیق شرکت داشته‌اند از آن دسته از افرادی نباشند که درمان آن‌ها موثر باشد. بدین معنی که در مواردی خاص نتایج، قابل عمومیت دادن نمی‌باشد و شواهد قوی یا تئوری‌هایی محکم وجود دارد که نشان دهد دسته‌ای از بیماران به صورت متفاوتی به درمان پاسخ می‌دهند.^(۱۸) به عنوان مثال داروهای *Blocker* نسبت به داروهای *Diuretic* کمتر بر روی سیاهپوستان نسبت به سفیدها در پایین آوردن فشار خون تأثیر دارد.^(۱۸) عوامل بیولوژیکی، اقتصادی، روانی، ژنتیکی و غیره می‌توانند در این ناهمگونی اثرات درمان موثر باشند و باید در نظر گرفته شوند.

تأثیر یک روش درمانی به خاطر اینکه جمعیت مورد درمان از میزان *Morbidity* یا ریسک مرگ یکسانی برخوردار نباشند نیز می‌تواند دچار اختلال شود. در اندازه‌گیری تأثیر درمان، بیماریارانی که در خطر بالاتری هستند (*High risk*) به طور معمول از درمان بیشتر تأثیر می‌گیرند.^(۱۸و۱۹) به عنوان مثال بیماریارانی که در ریسک بالای مرگ به دلیل ابتلاء به بیماری کرونر قلب (*CHD*) هستند با داروهای کاهنده کلسترول خطر مرگشان به مراتب بیشتر کاهش می‌یابد تا بیماریارانی که در ریسک پایین بوده‌اند. شاید بتوان مساله را این طور مطرح کرد که ۳۰ بیمار با خطر مرگ بالا، باید ۵ سال درمان شوند تا جان یکی از آنها نجات پیدا کنند. در حالیکه ۳۰۰ بیمار با ریسک پایین مرگ باید در طی ۵ سال درمان شوند تا فقط یکی نجات یابد.^(۲۰) بنابراین درمانی که برای یک بیمار با ریسک بالا ارزش بکارگیری دارد، ممکن است که ارزش امتحان کردن در بیماران با ریسک کم را نداشته باشد.

از دیگر مواردی که در حین بکارگیری یک روش درمانی در یک جمعیت مورد مطالعه باید توجه داشت شامل موارد عدم درمان (*Contraindication*) و عوامل اجتماعی فرهنگی جمعیت مورد مطالعه و اختلاف آن با بیماریارانی که با آنها سروکار داریم می‌باشد.^(۱۸)

سؤال ۳: آیا کسانی که در تحقیق انتخاب شده‌اند در تجزیه

و تحلیل و نتیجه‌گیری انتهائی شرکت داشته‌اند؟

حضور تمامی افرادی که ابتدا در تحقیق شرکت کرده‌اند در انتهای تحقیق و *Following* از جمله مواردی است که به قدرت *RCT* می‌افزاید. این کار از ایجاد انحراف در نمونه‌ها (*Subject bias*) جلوگیری می‌کند. اصولاً در طی یک تحقیق نمونه‌ها به دلایلی مختلف ممکن است از ادامه تحقیق صرف نظر کنند. مواردی چون نقل مکان از شهر، از دست دادن علاقه، اثرات سوء ناشی از درمان یا مرگ می‌تواند در بیرون رفتن نمونه‌ها از تحقیق نقش داشته باشد. بیرون رفتن نمونه‌ها از تحقیق ممکن است باعث شود آن دسته از بیماران که احتمالاً نتایج مثبتی از درمان می‌گرفته‌اند در نتیجه‌گیری انتهائی حذف و بنابراین باعث انحراف (*bias*) در مقایسه گروه‌ها شود (یا بر عکس).^(۱۰) همچنین این روند منجر به کاهش قدرت (*Power*) تحقیق می‌شود و حتی ممکن است منجر به عدم تعادل در بیماران موجود در گروه شود.^(۱۷) بنابراین محاسبات آماری حذف شدن نمونه‌ها بیشتر از ۱۰ درصد در یک تحقیق کوتاه مدت و بیشتر از ۱۵ درصد در یک تحقیق بلند مدت نتایج تحقیق را بی اعتبار می‌کند.^(۲)

سؤال ۴: آیا کلیه کسانی که در تحقیق شرکت داشته‌اند (نمونه و آزمونگر) نسبت به درمان بی اطلاع بوده‌اند؟

خطای آزمونگر (*Investigator bias*) ممکن است به هنگام مشاهده رخ دهد. بدین صورت که بدون قصد ممکن است نتایج غیردلخواه را در جهت تأیید فرضیه خود حذف کرده یا کنار بگذارد. از این انحراف می‌توان با فرایند بی اطلاع قرار دادن (*Blinding*) جلوگیری کرد. بهترین روش انجام این کار به صورت (*Double blinded*) می‌باشد. بی اطلاع گذاشتن بیمار از طریق انتخاب گروه *Placebo* انجام می‌گیرد. آزمونگرها هم از طریق بی اطلاع بودن از گروه درمان و گروه *Placebo* می‌توانند از خطا مصون بمانند. در مواردی که بی اطلاع گذاشتن بیمار و محقق امکان پذیر نباشد حداقل باید آزمونگرها بی اطلاع باشند (*Single-blinded design*)^(۱۷)

سؤال ۵: آیا گروه درمان و کنترل (شاهد) در شروع

کارآزمایی یکسان بوده‌اند؟

ما انتظار داریم در طی فرایند تصادفی سازی (در صورت کافی بودن حجم نمونه) گروهها در شروع کارآزمایی از خصوصیات یکسان برخوردار باشند. به عبارت دیگر تنها اختلاف دو گروه در نوع درمان باشد. از آنجا که همیشه این کار عملی نیست اغلب محققین یک گروه مقایسه انتخاب می‌کنند. این اطلاعات در اولین جداول مقاله ارائه می‌شود. عوامل قابل اندازه‌گیری که در ابتدای کار باید مورد توجه باشند عبارتست از: خصوصیات بیمار (سن، جنس، قد، شدت بیماری، یافته‌های آزمایشگاهی، وزن، وضعیت اقتصادی - اجتماعی) و سایر موارد قابل اندازه‌گیری مانند قدرت عضلات، دامنه مفصل، تست ورزش، $V O_{2max}$ و غیره می‌باشد. باید توجه داشت که موارد فوق جزء خصوصیات شناخته شده می‌باشد و احتمال دارد خصوصیات ناشناخته‌ای وجود داشته باشد که ما از آن بی اطلاع باشیم.^(۱۰)

در صورتی که اختلاف در شروع تحقیق زیاد باشد ممکن است اعتبار تحقیق را زیر سؤال ببرد اما همه چیز از دست نمی‌رود چرا که روشهای آماری وجود دارد تا اختلاف موجود در خصوصیات را تنظیم کنند.^(۱۱) در چنین مواردی باید دقت شود که آیا محقق به مورد فوق اشاره کرده و آن را تنظیم کرده است یا نه؟

سؤال ۶: آیا گروهها به صورت مساوی درمان شده‌اند؟
با گروه درمان و گروه شاهد باید بطور مساوی رفتار شود مگر در نوع درمانی که به آنها ارائه می‌گردد. محققین در کارآزمایی‌های تجربی اغلب تمایل دارند که وقت و توجه بیشتری به گروه درمان نسبت به گروه شاهد نشان دهند. این ممکن است باعث تأثیر عوامل دیگری به غیر از درمان در نتایج گردد.^(۲)

سؤال ۷: آیا نتایج به همان اندازه که از نظر آماری معنی دار هستند از نظر کلینیکی هم اهمیت دارند؟

آمار، تکنیک ریاضی جمع آوری، منظم کردن، توضیح دادن، و تحلیل و تفسیر داده‌های عددی می‌باشد.^(۲) منظور از استفاده از آمار در تحقیق، متقاعد کردن خواننده به معتبر

بودن نتایج تحقیق است.

اعتبار به دو صورت در آمار مورد نظر است:

- اعتبار داخلی (*Internal Validity*) نشان دهنده خوب بودن تحقیق می‌باشد و تعیین می‌کند آیا نتایج منعکس کننده حقیقت هستند یا یک رخداد تصادفی می‌باشند.

- اعتبار خارجی (*External Validity*) بررسی می‌کند که آیا نتایج قابل تعمیم به بیماران خارج از تحقیق می‌باشند.

در محاسبه‌های آماری وقتی که نتایج را با واقعیت مقایسه می‌کنیم دو نوع اشتباه *TYPE I* و *TYPE II* ممکن است رخ دهد. در اشتباه *TYPE I* محققین یک اختلاف معنی داری بین گروهها پیدا می‌کنند در حالیکه در واقعیت چنین اختلافی موجود نیست. متأسفانه محقق نمیدانند کی و چگونه این اشتباه رخ می‌دهد و فقط می‌توانند حدس بزنند احتمال وقوع آن چقدر می‌باشد.^(۲) معمولاً در محاسبه‌های آماری، آلفا، حداکثر ریسک قابل قبول برای اشتباه *I* *TYPE* را کمتر از ۵٪ در نظر می‌گیرند تا حداقل امکان وقوع اشتباه *I* *TYPE* وجود داشته باشد.^(۲۱) محققین اغلب معنی دار بودن نتایج آزمون را مورد توجه قرار می‌دهند و صرفاً رسیدن عدد *P* (احتمال تصادفی بودن نتایج) به کمتر از آلفا یا $P < ۰.۰۵$ آنها را متقاعد می‌کند و اهمیت کلینیکی نتایج فراموش می‌گردد.^(۲۱) اهمیت کلینیکی به بزرگی اختلاف در نتایج بین گروه درمان و شاهد اطلاق می‌گردد. در یک تحقیق بزرگ (با تعداد زیاد نمونه) اختلاف بسیار کم بین دو گروه می‌تواند معنی دار (از نظر آماری) گزارش شود در حالی که از نظر کلینیکی از اهمیت قابل توجهی برخوردار نمی‌باشد. به عنوان مثال در یک تحقیق که اثر دو داروی ضد فشار خون^(۲۲) را روی ۱۰۰۰ بیمار بررسی می‌کند، اختلاف حدود ۳ mmHg را از نظر آماری نشان می‌دهد در حالیکه از نظر کلینیکی این نتیجه از اهمیت جزئی برخوردار می‌باشد.^(۲)

باید توجه داشت که کوچک بودن *P* کمتر از ۵٪ فقط نشاندهنده احتمال کمتر در وقوع اشتباه *TYPE I* می‌باشد. نتایج آزمون آماری نباید ما را اغواء کند بطوریکه نتایج را چشم بسته قبول کنیم بلکه باید اهمیت کلینیکی آن را هم

در نظر بگیریم. (۲)

می دهد که یک تحقیق هیچ اختلافی بین گروهها نشان ندهد در حالیکه این اختلاف به واقع موجود می باشد. مقادیر احتمال رسیدن به چنین وضعیتی را در ۲ یا ۲۰ درصد تنظیم می کنند. توان آزمون (*Power*) توانائی پیدا کردن این اختلاف را دارا است (در صورتیکه در واقعیت موجود باشد) و به صورت *I-B* محاسبه می گردد.

در مطالعاتی که دارای نتایج منفی می باشند، محاسبه توان می تواند بررسی کند که آیا حجم نمونه به کار رفته در تحقیق کافی بوده است یا به علت تعداد کم، محاسبه آماری نتوانسته است یک تفاوت معنی دار را نشان دهد. معمولاً توصیه می شود که محقق قبل از شروع یک کار تحقیقی، محاسبه توان را انجام دهد تا تعداد نمونه مورد نیاز را تعیین کند. در واقع فقط ۳۲٪ از کار آزمائی های تصادفی دارای نتایج منفی که در بین سالهای ۱۹۷۵ تا ۱۹۹۰ در مجلات منتشر شده اند دارای محاسبه حجم نمونه بوده اند. (۳) نگاه دیگر در همین مقالات نشان می دهد اکثریت این تحقیقها از حجم نمونه به مراتب کمتری از آنچه مورد انتظار است برخوردار بوده اند که منجر به توان آماری پایین گردیده است. (۴) بنابراین هنگام مطالعه یک مقاله که نتایج منفی دارد، از خود پرسید که آیا حجم نمونه کافی بوده است و در صورتی که در تحقیق مورد بحث محاسبه توان انجام نشده باشد، به نتیجه گیری آن با تردید نگاه خواهیم کرد. سوال ۹: آیا می توان نتایج درمانی جدید را در مورد بیماران خود به کار ببریم؟

به هنگام مطالعه یک مقاله در مورد درمانی جدید باید به شرایط انتخاب نمونه ها و شرایط کنار گذاشتن نمونه ها توجه شود. در صورتیکه بیماران شما با شرایط فوق مناسبت دارند شکی نمی ماند که نتایج فوق را در مورد بیماران خود می توان به کار بست. در صورتیکه شرایط مطابقت نداشت باید از خود پرسیم آیا دلیل قابل قبولی وجود دارد که درمان جدید را در مورد بیماران خود بکار نبریم. موارد اختلافی چون جنس، سن، شدت بیماری، منطقه جغرافیایی، انتخاب بیماران، سودمندی درمان نسبت به ضررهای آن، در نظر گرفتن میزان ریسک مرگ بیمار و غیره

در سالهای اخیر محققین بصورت فزاینده ای از فاصله اطمینان (*Confidence interval*) به جای عدد *P* استفاده می کنند. اشکالی که در ارتباط با عدد *P* وجود دارد این است که هیچ اطلاعی در مورد اندازه اختلاف (*Size of Difference*) یا ارتباط بدست آمده نمی دهد. همچنین عدد *P* فقط نتیجه را بصورت سیاه و سفید ارائه می دهد. یعنی یا نتایج معنی دار هستند یا نیستند. در طرف مقابل، *CI* یک دامنه عددی را به وجود می آورد که با درجه احتمال بالا شامل مقادیر حقیقی آزمون می باشد و می تواند اطلاعات بیشتری از عدد *P* تنها بدست دهد. (۲۲) مقدار استاندارد برای فاصله اطمینان ۹۵ درصد می باشد (که محدوده ای است با ۹۵ درصد احتمال در برداشتن ارزش حقیقی). هرچقدر تعداد نمونه ها بیشتر باشد، فاصله اطمینان (*CI*) دقیقتر و دامنه آن محدودتر میگردد.

فاصله اطمینان شامل مقادیری مربوط به فرضیه صفر *Null Hypothesis* می باشد. همانطور که می دانیم هدف نهائی فرضیه صفر رسیدن به مقادیر صفر می باشد. به این معنا که فرضیه صفر ادعا می کند که هیچ اختلافی بین میانگین گروه درمان و گروه شاهد وجود ندارد. بدین ترتیب اگر مقادیر فاصله اطمینان شامل عدد صفر باشد نتیجه گیری می کنیم که هیچ اختلاف معنی داری بین گروهها وجود ندارد و بر عکس. (۲۳) به عنوان مثال در یک *RCT* که داروی *A* و *B* را مقایسه می کند، ۵۰٪ بیماران با داروی *A* و ۴۶٪ با داروی *B* بهبود یافته اند. با استفاده از آزمون آماری و با در نظر گرفتن $P < 0.05$ اختلاف معنی داری بین دو درمان مشاهده میشود. در این حال فاصله اطمینان، ۹۵ درصد محاسبه شده و در دامنه عددی ۱۰-۰ میباشد. در مثال فوق با وجود اختلاف معنی داری که از نظر آماری مشاهده می گردد، اختلاف از نظر کلینیکی معنی دار نمی باشد.

سوال ۸: اگر نتایج یک کار آزمائی منفی باشد آیا توان آزمون محاسبه شده است؟

همانطور که ذکر شد انحراف *TYPE II* موقعی رخ

خیل زیاد مقالات تعداد کمی کامل خواهند بود. تجزیه و تحلیل منتقدانه بندرت سیاه و سفید می‌باشد بلکه اغلب رنگهای خاکستری با سایه روشن‌های مختلف را نشان می‌دهد.

فقط قضاوت آگاهانه شما می‌تواند تعیین کند تا چه درجه‌ای از خاکستری، قابل قبول برای به کارگیری درمان جدید می‌باشد. بکارگیری نکات، اصول و تکنیکهای ارائه شده در این مقاله می‌تواند شما را در تشخیص مطمئن‌تر شدت رنگ خاکستری مورد نیاز راهنمایی کند.

نیاز به قضاوت شمارا قبل از بکارگیری درمان جدید اجتناب‌ناپذیر می‌سازد. همچنین مواردی چون هزینه درمان جدید و در دسترس بودن آن می‌تواند در بکارگیری آن تاثیر عمده داشته باشد.^(۱۷،۲) علاوه بر موارد فوق عوامل اجتماعی و فرهنگی نیز می‌تواند در تعیین مناسب بودن یک درمان شرکت داشته باشد.

سخن آخر اینکه باتمیرین و به کارگیری یک *Check List*، شما می‌توانید یک ارزیابی منتقدانه روی یک مقاله را در مدت چند دقیقه انجام دهید. باید در نظر داشت که از

References

- 1- Mc kibbon K, hayword R, Haynes B. *The medical literature as a source. For evidence based care. ACP journal club.* <http://hiru.mcmaster.ca>.
- 2- Geyman P. *Evidence-based Medicine in primary care in: Geyman P, Deyor R. Evidence-based clinical practice. Butterworth ; 2000.ch.*
- 3- Stiell I, Wells G. *Multicentre trial to introduce the ottawa ankle rules for use of radiography in acute injuries. BMJ 1995;311:594-97.*
- 4- *Evidence-Based medicine working group. A new approach to teaching to practice medicine . Jama 1992;260:2420-25.*
- 5- Haynes RB, Sackettd. *Transferring evidence from research into Practice. ACP j Club 1996;125:A14-A16.*
- 6- Field MJ, Lohe Kn. *Guidelines for clinical Practice. Washington : national academy of science , Institute of medicine; 1992.P.321. Online national academy press reading room.*
- 7- Ellis J, Muliigan I, Rowe J, sackett DL. *Inpatient general medicine is evidence-based. Lancet 1995;346:407-410.*
- 8- Tsuruoka K, Koki M. *Evidence - based gerenal practice . Drug treatment in gerenal practice in Japan. British med j 1996;313:114-116.*
- 9- Nordin - johansson A, Asplundk . *Randomized Controlled trial and consensus as a basis for interventions in internal medicine . J int med 2000;247:94-104.*
- 10- Guyatt GH, Sackett DL. *Vsersi guides to the medical literature. How to use an article about therapy and prevention. JAMA 1993;270:2598-601.*
- 11- Fletcher R, Fletcher S. *Keeping clinically up-to-date. Evidence-based approaching to the medical literature. Gen intern Med 1997;121:60-61.*

- 12- Lock S. Demaon editorial peer review work. *Ann Int Med* 1994;60-61.
- 13- Begg C, Cho M, East Wood S. Improving the quality of reporting . *JAMA* 1996;276:637-39.
- 14- Justice A, Berlin J. Do readers and peer reviewers agree on manuscript quality? *JAMA* 1994;272:117-19.
- 15- Greenhalgh T. How to read a paper. *BMJ* 1997;315:243-6.
- 16- Franks P. Clinical trial . *Fam Med* 1988;20:443-8.
- 17- Helewa A, Walker JM. *Critical evaluation of research in physical Rehabilitation*. WB Saunders co ;2000.P.83.
- 18- Sheldon TA, Guyatt GH. When to act on the evidence. *BMJ* 1998;317:139-42.
- 19- Davey Smith G, Eggar M. Who benefits form medical intervention? *BNJ* 1993;308:72-4.
- 20- Davey Smith G, Sons F, Sheldon TA. Cholestrol Lowering and mortality. The importance of Considering initial level of risk. *BMJ* 1993;306:1367-73.
- 21- Greenhalgh T. How to read a paper. *Statistic for non-statistician* . *BMJ* 1997;315:346-66.
- 22- Braitman L. Confidence interval assess both clinical and ststistic Significant. *Ann int Med* 1991;114:515-17.
- 23- Gardner MJ, Altman DG. Confidence interval rather than p value. Estimation rather than hypothesis testing. *BMJ* 1986;292:746-50.
- 24- Moher D, Dulberg. Statistical power, Sample size, and their reporting in RCT. *JAMA* 1992;272:122-24.