التقييم النقدي للتجارب السريرية

02 د. يوسف لطيفة

محتوى مجاني غير مخصص للبيع التجاري

الطب المسند بالدليل | EBM

السلام عليكم

نصل وإيّاكم للمحاضرة الثانية بقسم الدكتور يوسف لطيفة والأخيرة بمادة الطب المسند بالدليل EBM، والتي سنتحدث فيها عن التقييم النقدي للتجارب السريرية، لنبدأ...

مقدمة

- القاعدة الأساسية في قراءة الدراسات والتجارب السريرية: "لا تصدق كلّ ما تسمع، بل قيّمه تقييماً ناقداً"، فوجود الأدبيات المنشورة في مقال لا يعني أنها كلها جيدة.
- تعدّ مجلة The New England Journal of Medicine، ومجلة The Lancet رقم واحد طبياً في العالم حتى الآن، وتعدّ الأبحاث التي تُنشر فيهما من أرقى الأبحاث.
- حتى وإن كان البحث منشور في هذه المجلات يجب أن نقرأه قراءة نقدية، لأن هناك احتمال للخطأ دائماً.[1]

أهمية التقييم النقدي

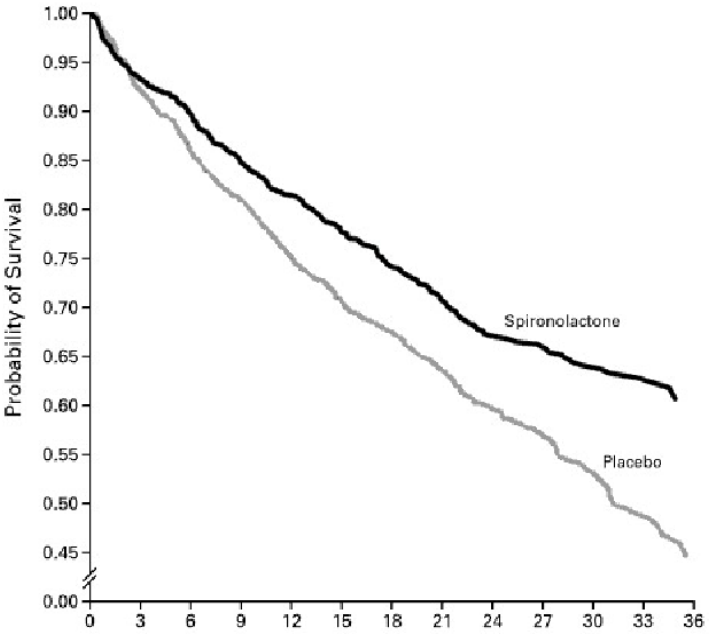
- عند الممارسة السريرية من الممكن أن نحتاج لمعلومة من بحث ما أو أن تقدم لنا شركات الأدوية دراسة حول دواء جديد لمرض معين، فتأتي شركة الأدوية وتعرض علينا هذا الدواء على أنه أفضل دواء للمرض ويكون مرتفع الثمن.
- ولإثبات ذلك تُقدم الشركة لنا دراسات تثبت فعالية هذا الدواء لكن لا يجب أن نصدق هذه الدراسات مباشرةً، وإنّما نقوم بتحليل وتقييم نقدي لها للتأكد من مصداقيتها (إلا في حال كان الدواء قديم ومعروف فلا داعي للقيام بتقييم نقدي).

alkamal
22330507

[1] ذكر الدكتور مثال عن أحد الأبحاث المنشورة في NEJM حول كورونا الذي تم حذفه لاحقاً لأنه تبين أنه مزوّر.

◈ وللقيام بذلك يجب أن نطلب النص الكامل للدراسة وليس الملخص ونقوم بنقده، فهنا تأتي الأهمية الكبيرة للتقييم النقدي.

مثال عن أهمية التقييم النقدي:



❋ نشرت مجلة NEJM عام 1999 مقال عن علاقة السبيرونولاكتون (دواء مدر للبول) بالمِراضة وحدوث الوفاة لدى مرضى القصور القلبي المزمن.

❋ أُخذت عينة مناسبة لإجراء RCT عليها، وقسّمت لمجموعتين:

- المجموعة A: تم إعطاءهم السبيرونولاكتون.
- المجموعة B: تم إعطاءهم Placebo.

❋ فكانت النتيجة: أن الدواء يقلل المراضة والوفاة مقارنة مع الـ Placebo.

❋ معلومات عن البحث:

- الجمهرة Population: NY class IV[2]، LVEF[3] <35%
- التداخل Intervention:
  - ✓ السبيرونولاكتون Spironolacton 25-50 mg/d.
  - ✓ الدواء الغُفل Placebo.
  - ✓ العشوائية Randomization، التعمية Blinding، تحليل نية العلاج ITT analysis.
- النتيجة (الحاصل) Outcomes:
  - ✓ أُجريت الدراسة لمدة سنتين وتم دراسة معدل الوفاة خلالهما.
  - ✓ معدل الوفاة خلال سنتين للخاضعين للتجربة (تم إعطائهم الدواء) EER[4]=35%.
  - ✓ معدل الوفاة خلال سنتين للشواهد (تم إعطائهم Placebo) CER[5]=46%.

---

2 تصنيف جمعية أطباء القلب في نيويورك New York Heart Association لقصور القلب.

3 Left Ventricular Ejection Fraction كمية الدم المقذوف من البطين الأيسر في الانقباضة الواحدة، وهو مقياس لقوة ضخ القلب، النسبة الطبيعية له أكبر من 55 بالمئة.

4 Experimental Event Rate.

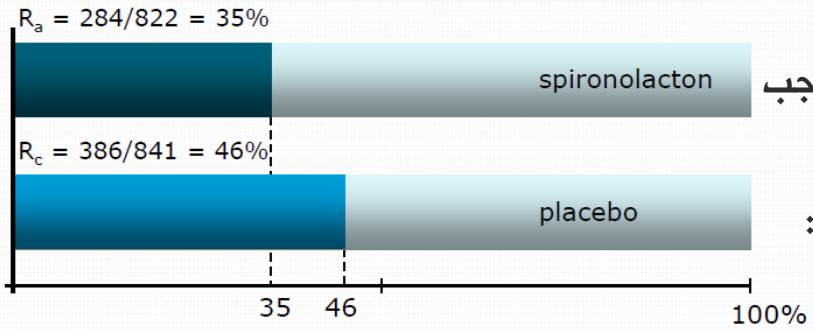
5 Control Event Rate.

- ✓ والفرق بينهما $p < 0,001$, (7–16) 11% = (95% CI) RD[6].
- ✓ فرضية العدم H0: بفرض عدم وجود فرق: ما هو احتمال العثور على فرق يساوي أو أكبر من 11% بالصدفة؟

إنقاص الخطر المطلق[7] ARR = الخطر في مجموعة الشاهد (المقارنة) - الخطر في مجموعة المعالجة:

	Death	
	Yes	No	total
spironolacton	284	538	822
placebo	386	455	841

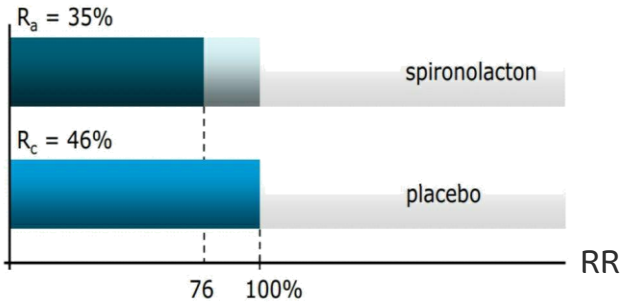
- ✓ ARR = $R_c - R_a$ = 46% - 35% =11%
- ✓ أي هناك 11 وفاة أقل عند معالجة 100 مريض.
- ✓ وهنا نقوم بطرح السؤال: كم مريض يجب أن نعالج لنتجنب حدوث وفاة واحدة؟



فنقوم بحساب **العدد اللازم للعلاج** NNT:

- ✓ NNT= 1/ARR = 1/0.11 = 9
- ✓ أي: 9 مرضى يجب أن يتعالجوا لمدة سنتين كل يوم حبة من الدواء لكي نتجنب حدوث وفاة واحدة.

إنقاص الخطر النسبي[8] RRR:



- ✓ يمثل نسبة الخطر التي يزيلها العلاج المستخدم.
- ✓ RRR= 1 - RR = 1 - 0.76 = 0.24

تحويل العلم إلى نقود: هو التطبيق العملي المادي لكل ما سبق.

- ✓ لدينا 9 مرضى، تتم معالجتهم لمدة سنتين، حبة كل يوم لكل مريض، أي:
  365 يوم × 9 مرضى × 2 سنة= 6570 حبة ثمنها 1000$.
- ✓ وبناءً على النتيجة يتم اتخاذ القرار من قبل الطبيب والمريض، أو إدارة المشفى، أو أهل المريض.
- ✓ هذا الدواء يعتبر رخيص فيكون من مسؤولية المشفى أو وزارة الصحة معالجة المريض.

[6] فرق الخطر Risk Difference.

[7] Absolute Risk Reduction.

[8] Relative Risk Reduction.

فريق التجارب السريرية The Trial Team

- أي دراسة أو تجربة سريرية أو تجربة RCTs يتم إجراؤها، يجب أن تتم ضمن فريق فيه مدير وباحثين ومساعدين، وأشخاص يقومون بتوزيع الاستبيانات، وآخرين يقومون بإدخال البيانات بعد انتهاء الاستبيانات على البرامج الإحصائية مثل Excel وSPSS[9]، ولهذا يجب وجود فريق كامل يضع خطوات العمل ثم يقوم به.[10]
- ويتكون فريق التجارب السريرية من العديد من الأفراد، وهم:

1. الباحث الأساسي Principal Investigator.
2. منسق (مدير) التجارب Trial Coordinator Or Manager.
3. مبرمج التجارب Trial Programmer.
4. مدير تنظيم المعلومات Data Manager.
5. سكرتارية التجارب Trial secretary.
6. إحصائي التجارب Trial Statistician لمراحل البحث المختلفة، وهي:
   - التخطيط Planning Phase.
   - التحليل المؤقت Interim Analysis.
   - التحليل النهائي Final Analysis.

بعد أن تحدثنا عن التقييم النقدي وأهميته، وعن أفراد فريق التجارب السريرية، ننتقل الآن للحديث عن خطوات التقييم النقدي...

[9] إثراء: Statistical Package For Social Sciences، وهو من البرمجيات الشائعة وسهلة الاستخدام في مجال البحوث الاجتماعية والصحية، حيث يقوم الباحث بإدخال بيانات بحثه إلى هذا البرنامج ويقوم البرنامج بقراءة البيانات وتحليلها واستخلاص النتائج من خلال مجموعة من العمليات الإحصائية الموجودة فيه.

[10] أرشيف.

خطوات التقييم النقدي

❖ سنقتصر على الحديث عن خطوات التقييم **النقدي للدراسات التجريبية العشوائية** RCTs دون التطرق لباقي تصاميم الدراسات، ويتم التقييم لهذه الدراسات ضمن هذه المعايير السبعة التالية:

1. تقييم حجم العينة Sample Size.
2. اتباع العشوائية (التعشية) Randomization.
3. القيام بالتعمية (الإخفاء) Blinding (Masking).
4. تقييم النتائج (الحواصل أو المخرجات) Outcomes.
5. تحليل نية العلاج Intention To Treat Analysis (ITT).
6. قياس فعالية العلاج Measurement of Treatment Effect.
7. تطبيق النتائج على المرضى Applicability (تعميم النتيجة).

❖ وتَوَفر هذه المعايير السبعة كلّها يقودنا إلى **مصداقية البحث السريري**.

ملاحظات هامة: (أرشيف)

- كل معيار من المعايير السابقة له طرق تنقده بشكل صحيح.
- أهم مجال من المجالات السابقة هو العشوائية، فعند إرسال أي بحث أو استبيان أو RCTS لأي مجلة فإن المقيّم سينظر أولاً للعشوائية فإذا كانت معتمَدة سيكمل البحث وإذا لم تكن معتمدة سيرفضه.
- فالعشوائية هي الأساس بأي بحث، فعند عدم اعتمادها سيصبح لدينا انحياز جذري، وبالتالي البحث ليس له أي قيمة.

سندرس الآن كل معيار على حدة...

حجم العينة Sample Size

العوامل الأساسية في اختيار العيّنة

✧ يجب أن يكون حجم العينة ممثلاً للمجتمع ليكون هنالك قيمة إحصائية للدراسة.

✧ يجب أن يتبع الباحث طريقة منهجية في تحديد حجم العينة.

✧ يجب أن نأخذ وضع المجتمع بعين الاعتبار كالتالي[11]:

⇦ إما أن يكون مجتمع مغلق: مثل مدرسة أو جامعة أي طلاب كلية الطب مثلاً.

⇦ أو يكون مجتمع مفتوح: مثل مدينة.

✧ بشكل عام لا تكون العينة كبيرة جداً، فمثلا في مجتمع حجمه 500 نحتاج لعينة تقدر بحوالي 235، وإن كان حجم المجتمع 10000 تكون العينة حوالي 415، إذاً العلاقة ليست طردية وإنما **لوغارتمية**.

الأمر الشائع بأنه كلما زاد حجم العينة كانت الدراسة أصدق أمر خاطئ.

✧ ولكن بنفس الوقت فإن إجراء دراسة على عدد قليل من المرضى غير مقبول ويُضعف من قيمة البحث مثل 30 أو 50 مريض، وهذا ما تتبعه بعض وسائل الإعلام مثل الإعلان عن نجاح الشوكولا في علاج الاكتئاب وذلك من خلال دراسة أُجريت على 50 شخص أو دراسة فعالية اليانسون في القضاء عل COVID-19 عند 50 مريض، وبالتالي هذه الدراسة تكون غير صحيحة وغير موثوقة إذ أن 50 مقابل مئات الآلاف والملايين يساوي صفر (أي ليس له قيمة).

✧ تجنب الانحياز باختيار المشاركين في البحث بأن يكون احتمال اختيار كل مشارك متساوي.[12]

✧ في أي دراسة **إحصائية** لأي متغير فإن الحد الأدنى لحجم العينة هو 60، فمثلاً إذا أردنا القيام بدراسة حول نسبة الرُهاب الاجتماعي عند طلاب الطب فإن عينة 35 شخص غير مقبولة.

✧ في أي دراسة **تجريبية** لأي متغير فإن الحد الأدنى لحجم العينة هو 30.

✧ هناك بعض الدراسات الإحصائية تتم على 30 أو 40 فرد، لكن ذلك في حالات قليلة جداً وعندها تكون النماذج الإحصائية ودراسات التحيز صعبة جداً.

✧ في حالة مرض نادر يُقبل أي حجم للعينة، لكن غالباً نمط الدراسة لن يكون RCTs.

✧ في حال مرض نسبة انتشاره معروفة فيجب أن يكون حجم العينة محدد بهذه النسبة، بينما في حالة مرض نسبة انتشاره غير معروفة، فإننا نضع رقماً تقريبياً أو المعادلة العامة.

✧ مثال: يمكن القيام بدراسة لمرض نادر، مثلاً داء هنتغتون على عينة من 35 شخص لكن حينها ستكون الدراسة وبائية إحصائية وليست RCTs، فيمكن أن تتم على 35 شخص كون المرض نادر وهي حالة خاصة ليس كما الحال العام.

[11] أرشيف.

[12] أرشيف لنهاية الصفحة.

ملاحظات: (أرشيف)

- الحد الأدنى لحجم العينة قريب للحد الأقصى مهما كان حجم المجتمع كبيراً، مثلاً مدينة عدد سكانها 50000 لن يكون حجم العينة الممثل لها 20000 بل سيكون بين 1000 - 1500 ويمكن التأكد من الرقم بدقة بواسطة البرامج الإحصائية.
- عندما يكون حجم المجتمع كبيراً وقادراً على إعطاء حجم عينة ممثلة له عندها يمكن إجراء المعادلات الإحصائية كاملة على العينة، بينما إذا كان عدد المجتمع قليلاً سيكون حجم العينة التي من المفترض أن تمثّله غير مناسب للبرامج الإحصائية مثل SPSS وعندها نلجأ لبرامج تخصصية واختصاصي بالإحصاء، للمساعدة بإكمال الدراسة.

المحاور التي يجب نقدها بالنسبة للعينة

أولاً: طرق حساب حجم العينة:

- لحساب حجم العينة يجب أولاً اختيار **المجتمع الممثل**، ثم حساب حجم العينة التي يجب أن اختيارها من المجتمع، ويوجد عدة طرق لحسابها منها:

1. الحساب بالقوانين الإحصائية:

- هو قانون قديم في علم الإحصاء، لم يعد يُستخدم حالياً.

2. الحساب الالكتروني:

- هو الأشيع عالمياً وأسهل من الحساب بالقوانين الإحصائية.
- يتوفر على الإنترنت حاسبات الكترونية (برنامج بسيط جداً)، نضع فيه نسبة مجال الثقة، ونضع حجم المجتمع العام ونختار Sample Size Calculater فيعطينا حجم العينة الممثلة للمجتمع التي تعطي مجال الثقة الذي حددناه.
- مثال: لإجراء دراسة عن انتشار الشقيقة عند طلاب الطب وعددهم 5000 (أي حجم المجتمع الممثل)، فنضع هذا الرقم مع مجال الثقة المطلوب 95% مثلاً فيعطينا البرنامج حجم العينة الممثلة لطلاب الطب في هذه السنة.
- يحتاج الحساب الالكتروني لوقت لا يتجاوز 10-15 ثانية لتطبيقه بشرط معرفة بعض المعلومات، فعند الحصول على معادلة الحساب من غوغل نكون أمام خيارين:
  - **إما أن تكون العملية بسيطة:** (استبيان مثلاً) عندها يتطلب معرفة حجم المجتمع الأساسي فقط.

- **أو تكون العملية معقدة:** (معرفة انتشار مرض ما مثلاً) عندها يتطلب معرفة حجم المجتمع الأساسي، نسبة الانتشار، عدد الذكور والإناث، مجال الثقة، ونسبة الخطأ.

✍ يمكن معرفة عامل انتشار دراسة ما أو مجال الثقة وغيرها من عدة مصادر، وهي: [13]

- إما أن نعرف نسبة الانتشار في البيئة نفسها من دراسات سابقة.
- وفي حال لم نجد فإننا نستعين بنسب البلدان المجاورة.
- وفي حال عدم توفرها من الممكن أن نستعين بنسب منظمة الصحة العالمية أو الاتحاد الأوروبي بشكل تقريبي، أو أي منظمة موثوقة تقوم بدراسات وبائية أو إحصائية.
- مثال: إجراء دراسة عن الرهاب الاجتماعي عند طلاب كلية الطب البشري، وكان لدينا معرفة من دراسة سابقة أنَّ نسبة انتشار الرهاب عند طلاب كلية طب الأسنان يساوي 15%، والتي تعتبر كلية (بيئة) مجاورة، فنعتمد بدراستنا أيضاً أنَّ نسبة انتشار الرهاب عند طلاب كلية الطب البشري تساوي 15%.

ثانياً: الأخطاء المُرتكبة:

♦ تتأثر صلاحية الدراسة Validity of Study بنوعين من الخطأ أو الانحياز:

1. خطأ منهجي Systematic Error:

✍ ناجم عن الخطأ من النمط الأول (α) Type I error.

✍ وعند التقييم يجب الانتباه إن كان الباحث نوّه على معالجة الأخطاء المنهجية الناتجة عن:

A. سوء القياس: الذي سيؤدي إلى انحياز المعلومات Information Bias.

B. عينة غير ممثلة للجمهرة: سينتج عنها انحياز الانتقاء Selection Bias:

- مثال: بحث أُجري على طلبة الطب فتم اختيار عينة من الذكور فقط، أو فئة أو فئتين.

C. تأثير محددات أخرى سيؤدي إلى خلط أو تتويه Confounding:

- مثال: أُجريت دراسة في السبعينات على أن شرب القهوة يسبب سرطان الرئة (هنا لم يأخذ الباحث بعين الاعتبار أن سرطان الرئة سببه التدخين وليس شرب القهوة[14]، إذاً شرب القهوة تعتبر في هذا المثال عامل **مُتوّه**).

[13] أرشيف.

[14] التدخين يترافق غالباً مع شرب القهوة.

2. خطأ الصدفة Chance Error:

- ناجم عن الخطأ من النمط الثاني Type II Error (β).
- لا بد منه، لأن الدراسة لا تشمل كامل الجمهرة.
- ينجم عن اختلاف الاعتيان Sampling Variability.
- تكرار الدراسة على عينة ثانية يعطي نتائج قريبة ولكنها مختلفة.
- يزداد كلما كانت العينة أصغر.
- لذلك فإن صحة النتائج هي أمر محتمل Probable، ويمكننا:
  - تحديد احتمال صحتها (القيمة P) بدراسة فرضية العدم Null Hypothesis.
  - قياس دقتها بحساب مجال الثقة Confidence Interval.

وسنفصل الآن في كلّ من فرضية العدم وقيمة P ومجال الثقة...

فرضية العدم:

- يجب تقييم كون الباحث وضع في فرضية العدم **عدم وجود فروق ذات قيمة إحصائية بين عينيتيّ البحث**، وأن الفرق الملاحظ هو نتيجة الصدفة (مثلاً: فرق الخطر = 0، نسبة الأرجحية=1).
- عند إعطاء دواء ما لمجموعة، وإعطاء الـ Placebo لمجموعة أخرى فيجب أن تكون النتائج متساوية ليكون الفرق 0 وتُقبل فرضية العدم، أما في حال كانت النتائج مختلفة نرفض فرضية العدم وهذا يعني وجود علاقة ذات قيمة إحصائية.
- نقيس احتمال Probability صحة هذه الفرضية بفحص إحصائي:
  - فحص عياري Parametric Test، مثل: T-Test إذا كان للنتائج توزع طبيعي.
  - فحص غير عياري Non Parametric Test، مثل: Wilkinson Test إذا لم يكن للنتائج توزع طبيعي.

قيمة P (P Value):[15]

- هي احتمال صحة فرضية العدم.
- إذا كانت صغيرة جداً، تُرفض فرضية العدم وتعتبر النتيجة هامة إحصائياً Statistically Significant.
- القيمة المتعارف عليها هي $p<0.05$.

[15] سيتم التطرق لها أيضاً في فقرة لاحقة.

مجال الثقة Confidence Interval:

- هو المجال المحيط بقيمة ما، ويتضمن قيم الجمهرة باحتمالية محددة (درجة معينة من الثقة) وهي عادة 95%.
- كلما ضاق المجال كلما ازداد القياس دقة.

مكونات حساب حجم العينة[16]

1. الخطأ من النمط الأول I (α):

- يمثل الإيجابية الكاذبة، أي أن الشخص سليم لكن الفحوص تظهر أنَّ لديه المرض.
- تبلغ قيمته 0.05.

2. الخطأ من النمط الثاني II (β):

- وهو يمثل السلبية الكاذبة، أي أنَّ الشخص مصاب لكن الفحوص تظهر أنَّه سليم.
- تبلغ قيمته 0.20.

لحساب قوة الدراسة (البحث) نستخدم القانون التالي: (1-β)[17].

حجم العينة مرتبط بنمط الخطأ إما β أو α.

3. معدل الحدوث في مجموعة الشاهد Control.

4. معدل الحدوث في مجموعة العلاج.

الاعتيان العشوائي Randomization

أهمية التعشية

- عند قراءة أي بحث ننظر في البداية إلى العشوائية، وفي حال عدم وجود عشوائية فيه فلا داعي لإكمال القراءة لأن البحث يكون غير موثوق.
- وعندما يقوم باحث بإرسال بحثه لنشره في مجلة ما، فأول ما ينظر له المحكمون هو العشوائية، وفي حال عدم توافرها لا يتم قبول البحث ولا نشره.

[16] أرشيف.

[17] سبب استخدام هذا القانون يعود لحسابات إحصائية معقدة لن نقوم بالتفصيل فيها.

- مما سبق نجد أنه يجب علينا أن نضع عبارة "**تم اختيار العينة عشوائياً**" في أي بحث نجريه، فهذه العبارة لوحدها ترفع مستوى البحث بشكل كبير.
- مثال: بحث يستهدف دواء معين لخفض السكر، ولم يذكر أنه اختار العينة عشوائياً، قد يكون اختار أشخاصاً بعمر الـ 40، أصحاء، وقيمة السكر لديهم 130-150 من المؤكد أن فعالية الدواء ستكون ممتازة، بينما عند اختيار العينة بشكل عشوائي من جميع فئات المجتمع فستُثبت كون الدواء فعال أم لا.
- يوجد شرطان رئيسيان للاعتيان العشوائي:
  1) تحديد قواعد التوزيع.
  2) اتباع نفس القواعد بدقة طوال الدراسة.

طرق الاعتيان العشوائي

- العشوائية البسيطة Simple Randomization.
- الجدول الانتقائي العشوائي Random Table: وهو أفضل الطرق ويعد رقم واحد عالمياً، لكنه يحتاج لخبرة جيدة.
- حظر العشوائية Block Randomization.
- التعشية الطبقية Stratified Randomization.
- طريقة التقليل Minimization Method.
- التوزع العشوائي غير المتكافئ Unequal Randomization.
- إخفاء التخصيص Allocation Concealment.

مما سبق نجد أنه إضافةً لعبارة "تم اختيار العينة عشوائياً" يجب أن نكتب "تم الاختيار العشوائي بطريقة الجدول أو الطريقة البسيطة" ... الخ.

سنشرح منها الجدول الانتقائي العشوائي وحظر العشوائية وإخفاء التخصيص...(أرشيف عدا فقرة إخفاء التخصيص)

الجدول الانتقائي العشوائي:

- هو مجموعة أرقام مصفوفة بجانب بعضها ضمن جدول له خط أفقي وخط شاقولي.
- إذا أردنا أخذ عينة مقدارها 350 طالب من طلاب السنة الثالثة في كلية الطب الذين يبلغ عددهم 1500 مثلاً.

❄ لأجل الاعتيان نأخذ تسلسل طلاب الطب (ترقيمهم) وعندها يمكن أن نعتمد على حجر نرد نرميه مرتين لنعرف من أين نبدأ بجدول الانتقاء العشوائي.

❄ فمثلاً أظهر حجر النرد النتيجة 2،6 فنختار الرقم من جدول الانتقاء العشوائي الذي يكون إحداثيته 2،6 (إحداثية البدء)، ثم نكمل لنهاية الجدول وهذه الطريقة صحيحة 100% للاعتيان العشوائي.

❄ يمكن تطبيق طريقة الطرة والنقش بدل حجر النرد لمعرفة من أين نبدأ بالجدول، كما أنَّها مقبولة للاعتيان العشوائي لكنها ليست ذات قوة كبيرة، وتطبق لمعرفة أين نبدأ في جدول الانتقاء العشوائي.

حظر العشوائية Block Randomization:

❄ مثلاً في حال تمت دراسة على فعالية دواء ما في مشفى المواساة، وقامت الشركة المصنعة للدواء بالترويج أنَّ الطبيب الفلاني (طبيب مشهور) سيكون متواجداً في المشفى في اليوم الفلاني لعلاج المرضى، فعندها سيتوجه الكثير من المرضى له في هذا اليوم.

❄ بالمقابل تقوم الشركة باختراع أسباب لتقليل مرضى الحالات الشديدة في اليوم الذي يتم فيه اختبار فعالية الدواء، فيكون الهدف اختيار مرضى تكون درجة المرض لديهم خفيفة لكي يظهر الدواء نتائج وفعالية جيدة وهذا ما يسمى حظر العشوائية.

إخفاء التخصيص Allocation Concealment:

❄ نذكر أنَّه أحياناً تكون العشوائية حسب السحب، كالتالي:

- يكون لدينا مجموعة من المرضى نريد أن نوزعهم على مجموعتين وهما مجموعة العلاج ومجموعة الشاهد Placebo، فيتم توزيعهم من خلال مجموعة من المغلفات حيث يقوم كل مريض بسحب مغلف عشوائياً وداخل هذا المغلف إمَّا أن يكون مكتوب فيه مجموعة علاج أو شاهد، هذه العملية تدعى بإخفاء التخصيص، وتتم بشرطين وهما:

1. **مغلف معتم مختوم (تعمية المغلفات):**

- أي في حال تم وضع المغلف أمام ضوء ساطع فلا يمكن رؤية ما بداخله (نختار مغلفات بلون أسود والرقم أو الخيار المكتوب أيضاً بلون أسود كي لا تظهر على الضوء).
- ألا يعلم الشخص الذي يسحب المغلف لأي مجموعة ينتمي.

2. **التوزيع العشوائي عن بُعد:**
   - وهي طريقة مُفضلة.
   - حيث يتم إلغاء تكليف القائمين على توزيع المهمات، ويتم التوزيع عن طريق الهاتف أو الإنترنت بدلاً من ذلك.

يجب أن تتم التعشية بعيداً عن الأطباء القائمين على الدراسة من أجل تجنب التحيز.

طرق أخرى للاعتيان العشوائي

- هذه الطرق **غير مقبولة** أساساً للاعتيان العشوائي، ولكن في حال تم قبولها لصعوبة تطبيق طرق أخرى فإنها تُنقص كثيراً من قيمة الدراسة، وهي:
  - طريقة الطرة والنقش.
  - الاختيار حسب تاريخ الميلاد (حتى الأرقام كذا يصنفون في المجموعة A...).
  - الاختيار حسب رقم الدخول إلى المشفى (تاريخ القبول).
  - الاختيار حسب تاريخ رؤية المرضى في العيادة، كالتالي:
    - المرضى الذين يأتون إلى العيادة في الأسبوع الأول يصنفون في المجموعة A.
    - المرضى الذين يأتون في الأسبوع التالي يصنفون في المجموعة B.
  - طريقة يوم بيوم: مثلاً لأجل اختبار دواء ما على المرضى المراجعين، نطبق الدواء على مرضى أيام الأحد والثلاثاء والخميس وندخلهم للدراسة، بينما نترك مرضى السبت والاثنين والأربعاء خارج الدراسة:
    - هذه الطريقة غير مقبولة أبداً حيث من الممكن أن تأتي فئات من المجتمع بأيام معينة أكثر من بقية الأيام لسبب ما، يوم يوافق عطل مهن معينة، أو يكون التوجه نحو طبيب يداوم في يوم ولا يداوم في آخر.
    - وبالتالي فهي تناقض مبدأ العشوائية رغم أنها كانت تستخدم سابقاً.
- إن مشاكل هذه الطرق تأتي من **انفتاح** نظام التخصيص Allocation System (أي أن المشارك يعلم لأي مجموعة ينتمي للشاهد أم للتجربة).

- يوجد طريقة اسمها كسر الاعتيان تشبه بالمبدأ طريقة يوم بيوم لكنها مقبولة بال RCTs أحياناً، مثلاً لأخذ عينة من مرضى مراجعين في يوم ما نقوم بترقيم هؤلاء المرضى ثم نختار أرقام وفق طريقة انتقاء عشوائي لأخذ عينة منهم.

التعمية أو الإخفاء (وضع القناع) Blinding or Masking

التعريف:

❄ هي أن تجعل واحداً أو أكثر من الأشخاص الذين تتضمنهم التجربة غير واعين للتداخل الذي سوف يطبق عليهم، وذلك بهدف إنقاص خطر الانحياز الناجم عن الملاحظة قدر الإمكان.

❄ **ملاحظة:** التعمية الكاملة والممتازة هي التعمية التي تتم على المريض، والطبيب الفاحص، ومدخل البيانات، والمحلل الإحصائي.

الأنواع:

❄ تشمل التعمية ست درجات في الـ RCTs وهم:

1. المشاركون.
2. المسؤولون عن المقابلات (المدخلات)، مثل الأطباء.
3. المسؤولون عن العناية بالمشاركين، مثل السكرتارية.
4. المسؤولون عن تسجيل النتائج (المخرجات).
5. محللو البيانات.
6. المسؤولين عن كتابة نتائج الدراسة.

الدرجات:

✧ هناك خمس درجات للتعمية:

1. التعمية المفتوحة Open Lapel: كل شخص مدرك وواعي لمكانه من البحث.
2. التعمية الأحادية Single-Blind: المريض فقط لا يعرف الدواء المستخدم في الدراسة.
3. التعمية الثنائية Double-Blind: المريض والباحث (الطبيب) لا يعرفان الدواء المستخدم.
4. التعمية الثلاثية Triple-Blind: المريض والباحث والمسؤول عن تسجيل النتائج لا يعرفون الدواء المستخدم.
5. التعمية الرباعية Quadruple-Blind، ومن الممكن أن تزيد عن ذلك.

- ما يهمنا في التعمية هو عدد ودور الأشخاص الذين لم تتم تعميتهم وليس عدد المعميين.
- التعمية ليست ملائمة وممكنة دوماً، أي لا يمكن تطبيقها دوماً.

صورة طريفة لتوضيح الفرق بين التعمية Blinding والقناع Masking القناع قد يكذب وأقل مصداقية، بينما التعمية حقيقية وأكثر مصداقية.

تحليل النتائج Outcomes

- ◈ لا يتم قبول أي نتيجة نحصل عليها من الدراسة، لذلك يجب أن نقوم بتحليل النتائج لنعرف هل هي مقبولة أم لا.
- ◈ تصنف النتائج إلى ثلاثة أقسام رئيسية:

أولاً: النتائج الأولية (البدئية) Primary:

- ❄ وتقسم أيضاً إلى نوعين:

1. النتيجة الأولية الأولى One Primary Outcome:

- ✍ هي النتيجة الأكثر أهمية وتشمل الفائدة والضرر، ففي أي دراسة يتم إجراؤها ننظر أولاً إلى الفائدة أو الضرر الذي حصل نتيجة هذه الدراسة، وهل هذه النتائج متوافقة مع البحث أم لا.
- ✍ وهي المتغير الأكثر صلة للإجابة عن السؤال البحثي.[18]
- ✍ مثال (1): فعالية لقاح للـ COVID-19 بتقوية المناعة.
- ✍ مثال (2): السكتة الدماغية عند استئصال قسم من باطن الشريان السباتي، فهنا تكون السكتة الدماغية الضرر الأول لهذه العملية وبالتالي تكون **النتيجة الأولية الأولى**.

2. النتائج المركبة Composite Outcomes:

- ✍ **مثال**: الأدوية في مرض احتشاء العضلة القلبية MI، من الممكن أن تقودنا للنتائج التالية:
  - تسبب الوفاة أو تنقص احتمال الوفاة، أو تكون غير قاتلة.
  - تسبب الإقامة بالمشفى بسبب الإصابة بمتلازمة الشريان التاجي الحادة ACS.

[18] إثراء.

- يمكن أن تزيد أو تنقص مدة الاستشفاء.

- كما رأينا أنَّه يمكن أن يكون هنالك أكثر من نتيجة أولية لدواء واحد لذلك هذه النتائج من الممكن أن **تكون مُضلِلَة**.
- قيمة هذه النتائج تعتمد على مدى التشابه بين مقدار الخطر لدى المريض، وتأثيرات العلاج وعدد الحالات الأخرى المعقدة (المركبة).
- تُهمل النتائج إن وُجد تنوع كبير بين الحالات (تعقيدات متعددة).

ثانياً: النتائج البديلة Surrogate Outcomes:

- تنتج عن الحالات النادرة في الدراسات السريرية الهامة.
- نقوم بتسجيل النتائج البديلة الإيجابية لنشرها ببحث مرفق، أو من أجل استكمال أبحاث أخرى حول هذا الموضوع.
- أمثلة:
  - دراسات تخص الحماية الخلوية من مضادات الالتهاب غير الستيروئيدية NSAIDs.
  - دراسات عن القرحات التي تُرى بالمنظار مترافقة مع النزف أو القرحة المعدية المثقبة PU.

ثالثاً: النتائج الثانوية Secondary Outcomes:

- هي عبارة عن متغيرات مهمة للسؤال البحثي (مثل الـ SE للأدوية) يتم تقييمها أيضاً.
- وهي نتائج إضافية يتم رصدها للمساعدة في تفسير النتيجة الأولية.[19]
- يتم الانتباه والنظر للنتائج الثانوية إذا لم يكن هنالك تغيير في النتائج الأولية.
- في النتائج الثانوية يكون لدينا أيضاً فوائد ثانوية أو أضرار ثانوية.
- النتائج الثانوية تتبع لتصميم الدراسة.
- أمثلة:
  - لقاح الـ COVID-19 يرفع المناعة ضد فيروسات أخرى غير الكورونا.
  - المقارنة بين دواءين لمعالجة قصور القلب الاحتقاني التالي لاحتشاء العضلة القلبية، حيث أنَّ الدواء الأول يحصر مستقبلات الأنجيوتنسين، بينما الثاني يحصر الكالسيوم، وتم إعطاء كل دواء منهم لمجموعة من المشاركين في الدراسة، وبعد ذلك تمت دراسة النتائج التي نتجت عن كل دواء وكانت على نوعين:

[19] إثراء.

- نتائج إيجابية: مثل تحسين وظائف القلب.
- نتائج سلبية (التأثيرات الجانبية): مثل زيادة مدة المكوث في المشفى.
- إضافة لدراسة التكلفة المادية واضطرابات النظم القلبية، وكل ذلك يمكن أن يُدرَس إما ضمن النتائج الأولية أو الثانوية حسب الحالة.
- ومن الممكن أن نقترح نتائج بديلة عن هذه الأدوية.

تحليل نية العلاج Intention To Treat Analysis (ITT)

- عند إعطاء دواء معين لمجموعة من الأشخاص لدراسته تأثيره عليهم من الممكن لبعض المشاركين ألا يكملوا الدراسة بسبب:
  - الوفاة أو تأثيرات الدواء الجانبية عليهم.
  - التشخيص الخاطئ، أو الانسحاب بسبب عدم الالتزام.
- كل ما سبق يؤدي إلى حدوث **فقدان للمعلومات** في الدراسة.
- النسبة المسموحة للبيانات المفقودة يجب ألا تتجاوز 10% (وبالتالي تكون المتابعة أكبر أو تساوي 90%)، ونقوم بمعالجتها إحصائياً.
- هناك أفكار عالمية لكيفية التعامل مع البيانات المفقودة وكما ذكرنا سابقاً يوفر لنا كوكرن بعضاً منها للتعامل مع البيانات المفقودة بنية العلاج.[20]
- حتى الآن غير متفق عالمياً على آلية موحّدة لتحليل نية العلاج.
- التعامل مع البيانات المفقودة ومعالجتها يجب أن يُكتب في بروتوكول البحث من قبل الباحث نفسه، فمن الممكن أن يعتبرها الأسوأ أو الأفضل أو متوسطة.
- كما يجب أن يذكر الباحث سبب عدم استكمال الدراسة من قبل المرضى، وأكثر الأسباب شيوعاً هي التأثيرات الجانبية لبعض الأدوية.
- مما سبق نجد أنه عند تحليل الدراسة يجب أن نعود للبروتوكول ونرى هل جميع المشاركين أكملوا البحث أم لا، ونرى كيف تمت معالجة البيانات المفقودة في حال وُجد فقد فيها.
- عندما نستثني مريضاً من تحليل الدراسة، من الممكن ألا نكون متأكدين أن عوامل الإنذار في المجموعتين متشابهة، وهذا يقود إلى إمكانية حدوث انحياز، ولتقليه نقوم بتحليل نية العلاج.

[20] أرشيف.

- بعض شركات الأدوية تخبئ نسبة المفقودين فلا تذكرهم بل تعتبر أن نسبة الذين بقوا هم العينة الأساسية وبهذا تؤثر على النتائج.
- الأفضل أن نعتبر النتائج المفقودة هي الأسوأ لكن معظم شركات الأدوية تقدم اعتراض على هذا الموضوع، حيث تتدخل الأمور المادية دوماً ولا يوجد مثالية دائمة، فيتم علاج الموضوع بجعل نصف البيانات المفقودة للأسوأ وليس كلها، وهناك آخرين يأخذون بيانات الجزء الأول من العلاج (قبل فقدان البيانات) ويكمل بقية نتائج الدراسة اعتماداً عليه.
- استراتيجية تحليل الدراسات التجريبية العشوائية RCTs[21] تعتمد على أن جميع المشاركين في الدراسة (سواء كانوا في مجموعة العلاج أو في مجموعة الشاهد) يتم تحليلهم سوية بصفتهم يمثلون ذراع (مجموعة) في العلاج، أي سواءً تلقوا العلاج أم لا، وسواءً أكملوا العلاج أم لا.

وبالتالي المشاركين العشوائيين (كل المشاركين) = المشاركين الذين تم تحليلهم.

- **باختصار:** عندما يقوم الباحث بالبدء بأي دراسة يكون لديه مخطط لسير الدراسة في كل ذراع فيها وتوقع للنتائج في حال جرت الأمور بالشكل الأمثل، أي في حال أكمل جميع المرضى ولم تحدث أية مشاكل أخرى، وهذا ما يتم النظر له عند الحديث عن نية العلاج، أي لا نتحدث عما حصل في الواقع وإنما نتحدث عما كان من المفترض أن يحصل.

أمثلة:

- دراسة تتضمن 200 مريض، قام 180 منهم بإكمال الدراسة وغادر 20 مريض منهم، فهنا يجب أن نعلم لماذا ترك هؤلاء الـ 20 مريض (معرفة السبب هو أهم قياس في النقد):
  A. فإذا تركوا لأسباب مادية فهذا لا يؤثر على النتائج.
  B. أما إذا تركوا بسبب التأثيرات الجانبية الخطيرة للدواء أو بسبب الوفاة فهذا مهم ويجب الانتباه له، وخاصة إذا كان للتداخل علاقة بموت المرضى.
- هناك قصة شائعة من حوالي 10 سنوات عن بحث لعلاج أحد السرطانات:[22]
  - يتم العلاج من خلال دواء يُعطى بجرعات كبيرة مما يؤدي إلى تثبيط نقي العظام وبعدها تتم زراعة النقي وبالتالي نكون قد عالجنا السرطان ونتمكن فيما بعد من السيطرة على تثبيط النقي.

[21] أرشيف

[22] أرشيف

- أقيمت هذه الدراسة في حوالي 10 مراكز في سويسرا وألمانيا وأمريكا وجنوب إفريقيا، لكن بعد ذلك لم يتم استكمال الدراسة وتم إيقافها بعد 6 أشهر في كل المراكز، لأن معظم المرضى الذين طُبقت عليهم جرعات كبيرة من الدواء ماتوا ولم تتم السيطرة على تثبيط النقي لديهم ماعدا مركز جنوب إفريقيا صرّح بأنَّ النتائج لديه ممتازة وأنه سيكمل الدراسة.
- بعد ذلك تمت العودة لهذه الدراسة التي أجريت في مركز جنوب إفريقيا للتحقيق فيها فقام الباحث بالتهرب.
- ثم جاء أمر من منظمة الصحة العالمية لكشف نتائج البحث وكيف تم العمل فيه وعندها ذهب وفد إلى جنوب إفريقيا فلم يجدوا المركز أساساً واكتشفوا أنه تم تطبيق الدواء على مريضة واحدة وتوفت والبحث كله كان كاذباً.
- والغريب أنه بعد مدة استدعي هذا الباحث الذي زور النتائج ليقوم بأبحاث في أمريكا.

ملاحظات:

- العبرة: لا تصدق كل شيء وتعلم كيف تنقد الدراسات والتجارب لتصل إلى الحقيقة.
- يوجد أيضاً ما يسمى بعدم التجانس Heterogeneity وهو يعني وجود اختلاف كبير بين نتائج الدراسة عما كان متوقعاً حدوثه عند إجراء مراجعة منهجية.

قياس تأثيرات المعالجة[23] Measurement of Treatment Effect

- عند إجراء أي تداخل أو معالجة يجب مراقبة تأثيره.
- مثلاً عند إعطاء الأسبيرين كوقاية لمرضى لديهم جلطات دماغية أو سوابق حوادث وعائية دماغية، فيجب أن نراقب أثر هذه المعالجة عليهم مع الزمن (مثلاً بعد 5 سنوات أو 10 سنوات).
- وننظر ماذا حقق الأسبيرين أي نُقيّم فعالية العلاج وتأثيراته (تحسين البقيا – عدم حدوث نكس للجلطات الدماغية...)، وهذه التأثيرات يمكن أن تكون إيجابية أو سلبية (تأثيرات جانبية).
- لقياس تأثيرات المعالجة يجب أن نقوم بتحليل ما يلي:

1. القيمة الاحتمالية Probability Value (P Value).
2. الخطر النسبي Relative Risk (RR).
3. نسبة الاحتمالات Odds Ratio (OR).

[23] من هذه الفقرة لآخر المحاضرة تم إعطائهم بسرعة في المدرج فتم الاستعانة بالسلايدات والأرشيف.

4. **مسافة (فاصلة) الثقة** Confidence Intervals (CIs).
5. **العدد اللازم للمعالجة** Number Needed To Treat (NNT).

◈ **قياس تأثير المعالجة يتم على البيانات المتوفرة في الدراسة سواءً تم تحليلها خلال سير الدراسة (تحليل مؤقت) أو بعد انتهاء الدراسة.**

سنفصل بالحديث عن كلّ تعداد مما سبق...

القيمة الاحتمالية P Value:

❄ **تعني احتمال أن نلاحظ فرق بين مجموعتي العلاج والشاهد عن طريق <u>الصدفة</u>.**

❄ **لكي تكون الدراسة مقبولة يجب أن تكون** P Value **أقل أو تساوي 0.05، وفي أغلب الدراسات يتم اعتماد القيمة 0.05 (5٪).**

❄ **إذا كانت** P Value **<u>أكبر</u> من 0.05 <u>فليس لها قيمة إحصائية</u> ومن المحتمل أن يكون الاختلاف الملاحظ بين المجموعتين قد حدث عن طريق الصدفة.**

❄ **أما إذا كانت** P Value **<u>أقل</u> من 0.05 يكون لها <u>قيمة إحصائية وليست سريرية</u>، وغالباً لا يكون الاختلاف الملاحظ بين المجموعتين تم عن طريق الصدفة.**

❄ **إنَّ قيمة 0.05 تعبّر عن الخطأ في الدراسة، بمعنى آخر يمكن أن نقول أنَّ النتائج الأكيدة والصحيحة التي سنحصل عليها من الدراسة لا تقل عن 95% والنتائج المسموح بالخطأ فيها لا تزيد عن 5%.**

❄ **إذاً يمكن اعتبار ال** P Value **هو <u>احتمال الخطأ في الدراسة</u>، فقد يكون لدينا نتائج خاطئة سواء ناجمة عن إيجابية كاذبة أو سلبية كاذبة أو غير ذلك لكن يجب ألا تتجاوز 5%.**

❄ **وبمعنى آخر أيضاً يمكن أن نقول أنَّه بوجود احتمال خطأ في الدراسة يساوي 5% يمكننا تعميم نتائج الدراسة بنسبة 95% وهذا يعني أنها صحيحة إحصائياً بنسبة 95% ولكن في حال كانت الدراسة صحيحة إحصائياً ليس بالضرورة أن تكون صحيحة سريرياً.**

❄ **مثال1:** تم إجراء دراسة للمقارنة بين دواءين أحدهما غالي الثمن ووُجد إحصائياً أنَّه أفضل من الـ Placebo ولكن تأثيراته الجانبية كثيرة وتفوق الـ Placebo وبالتالي هذا الدواء فعال إحصائياً لكنه غير فعال سريرياً، إما بسبب تأثيراته الجانبية، أو بسبب غلائه، أو أنه يوجد خطأ إحصائي بسبب المتغيرات أو عوامل أخرى.

❄ **مثال2:** إجراء دراسة حول تأثير الثوم على خفض الضغط، كان عدد المشاركين بالدراسة 200 مريض، وتم إعطاء 100 منهم الثوم و100 آخرين لم يتم إعطاءهم شيء، فنلاحظ أنَّ الضغط قد انخفض لدى 60% ممن تناول الثوم، أما الـ 40% المتبقية منهم لم ينخفض لديهم الضغط.

- في تحليل الدراسة السابقة للحكم على فعالية الثوم نجد أنه لا يمكن الحكم على الدراسة من النسب المئوية فهي ليس لها قيمة ما لم يتم إدخالها هي وجميع مواصفات العينة إلى برامج إحصائية لتعطينا القيمة الإحصائية لها.
- وهذا ما تقوم به بعض وسائل الإعلام فلا يجب أن نصدقها قبل نقدها، لأن **أي نسبة مئوية دون دراسة إحصائية كاملة لا يوجد لها أي قيمة.**

الخطر النسبي Relative Risk:

❄ للتعرف على معنى الخطر النسبي RR يجب أن نتعرف قبله على معنى مفهوم الخطر.

❄ **مفهوم الخطر:**

- هو عدد المرضى الذين أتموا معايير الدراسة لنهايتها مقسومين على العدد الكلي للمرضى الداخلين في الدراسة.
- **مثال:** دراسة أجريت لاختبار فعالية صاد حيوي لمرضى الإسهال:

⇦ **مجموعة العلاج:**

- كان عدد المرضى 10 وتم إعطاء 4 منهم صاداً حيوياً.
- بالتالي يكون مقياس الخطر لديها 0.4 = 4/10.

⇦ **مجموعة المقارنة (الشاهد):**

- كان عددهم 10 وتم إعطاء 1 منهم الـ Placebo.
- بالتالي يكون مقياس الخطر لديها 0.1 = 1/10.

❄ **الخطر النسبي Relative Risk:**

- هو حاصل قسمة مقياس الخطر عند **مجموعة المعالجة** على مقياس الخطر عند **مجموعة المقارنة**.
- وهو بالمثال السابق: $RR = \frac{0.4}{0.1} = 4$.

الاحتمالات/الأرجحية ونسبة الاحتمالات Odds And Odds Ratio:

❄ **الاحتمالات/الأرجحية Odds:**

- هي عدد المرضى الذين أتموا معايير الدراسة لنهايتها مقسومين على عدد الذين لم يتموها في الدراسة نفسها.
- في المثال السابق تكون الاحتمالات تساوي:
  - للمجموعة الأولى (العلاج): 0.66 = 4/6.
  - للمجموعة الثانية (الشاهد): 0.11 = 1/9.

❄ **نسبة الاحتمالات Odds Ratio:**

- هي قسمة الاحتمالات عند المجموعة الأولى (المعالجة) على المجموعة الثانية (المقارنة).
- وهي تساوي في المثال السابق 6=0.66/0.11.

❄ **تفسير قيمة RR&OR:**

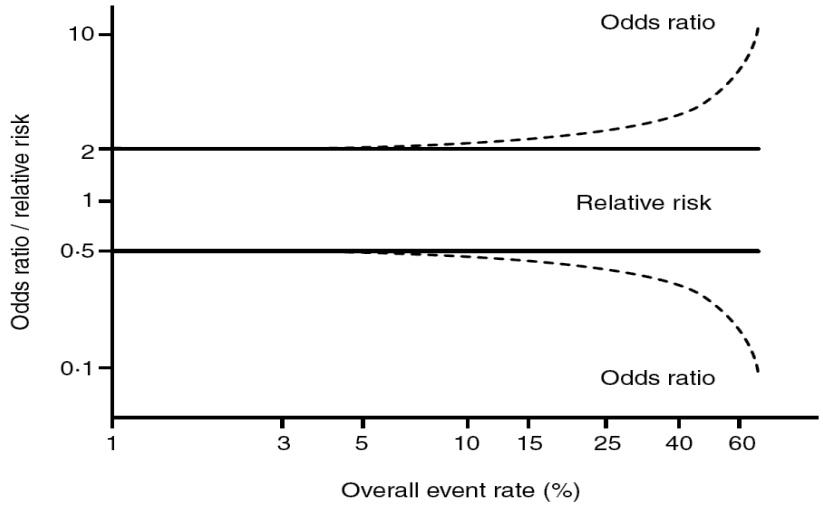
- يجب أن تتصاحب قيمة كل منهما مع قيمة مسافة الثقة CIs ونجد ثلاثة احتمالات لتفسير قيمتهما:
  - **RR or OR > 1:**
    - عندها يزداد احتمال أن تكون النتيجة في مجموعة العلاج[24].
  - **RR or OR < 1:**
    - ينقص من احتمالية أن تكون النتيجة في مجموعة العلاج.
  - **RR or OR = 1:**
    - ليس هناك اختلاف بالنتيجة بين مجموعة العلاج ومجموعة المقارنة.

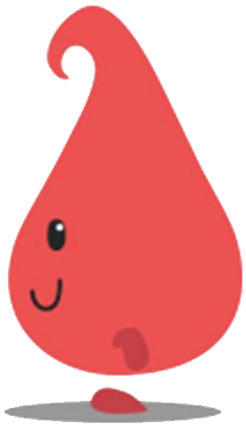
[24] توضيح: نعلم أن هذه النسبة هي عبارة عن ناتج قسمة المرضى في مجموعة العلاج على المرضى في مجموعة المقارنة، وبالتالي في حال كانت النسبة أكبر من 1 فإن ذلك يشير إلى أن البسط أكبر من المقام، لذلك يزداد احتمال أن تكون النتيجة في مجموعة العلاج.

❄ **مسار الـ OR&RR:**

بالمتابعة مع المخطط التالي نصل للنتيجتين التاليتين:

- إذا كانت نقطة النهاية تحصل بتكرار أقل من 15 يكون كلاً من RR وOR متقاربين.
- أما إذا كانت النتيجة شائعة أكثر أي زاد التكرار وزادت القيمة عن 15، سوف نجد أن OR تبتعد عن RR.





مسافة الثقة Confidence Intervals (CI):

❄ عندما نريد أن نختبر دواء جديد في عينة عشوائية من المرضى، فإن تأثير العلاج الذي نلاحظه في الدراسة يكون تقديراً حقيقياً وصحيحاً لتأثيره على الجمهرة كاملة التي أخذت منها العينة، وبالتالي فهو يحد من عدم الثقة، كونه يمثل مجالاً يحوي القيمة الحقيقية.

❄ إن نسبة 0.95 = 95% لفاصلة الثقة تشير إلى أننا متأكدين من صحة وفعالية العلاج بنسبة 95%.

❄ وهي تُعد **المقابل لمقياس الخطر ويتممان بعضهما البعض**، فمقياس خطر 5٪ يقابل فاصلة ثقة 95% ومقياس خطر 10% يقابل فاصلة ثقة 95% وهكذا.

❄ القيمة الأكثر استخداماً لمسافة الثقة هي 95%، كما يمكن أن يتم استخدام 90 و99% أحياناً لكن بشكل أقل.

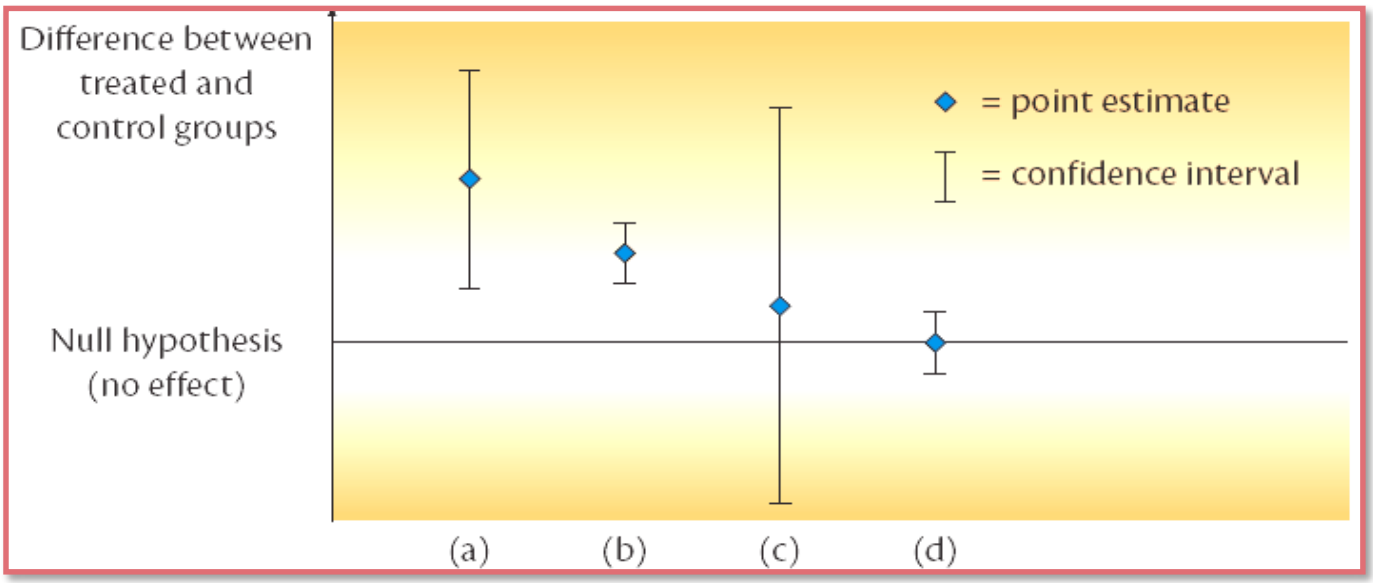
❄ كلما انخفضت قيمة مسافة الثقة كلما انخفضت قيمة الدراسة ومصداقيتها.

❄ عرض (نطاق) مسافة الثقة: إنَّ مجال مسافة الثقة **يدل على دقة القياس** في الدراسة، وكلما كان المجال أوسع كلما كانت الدقة أقل.

❄ إذا كان مجال الثقة يتضمن الرقم 1 فلا يكون هنالك اختلافات ذات دلالات إحصائية، أما في حال لم تتضمن الرقم 1 يكون هنالك اختلافات ذات دلالات إحصائية وهو المجال الأفضل.

الأهمية الإحصائية ومسافة الثقة Statistical Significance & CI:

❄ تابع الشرح مع المخطط لبيان أهمية مسافة الثقة:



❄ نلاحظ في الصورة وجود 4 خطوط طولية، كل خط منها يشير إلى دراسة معينة، أي أننا قمنا بإجراء 4 دراسات حول موضوع واحد (فعالية دواء مثلاً).

❄ يشير الخط الطولي إلى فاصلة الثقة، وكما ذكرنا منذ قليل أنه كلما كان مجال فاصلة الثقة أوسع (أي أن الفرق بين مجموعة العلاج ومجموعة الشاهد أكبر) كلما انخفضت دقة الدراسة لذلك نجد ما يلي:

- الدراسة B وD مرتفعة الدقة (لقصر مجال الثقة).
- الدراسة A وC منخفضة الدقة (لطول مجال الثقة).

❄ أما بالنسبة للشكل المعين في منتصف الخط الطولي لكل دراسة فهو يشير إلى القيمة الإحصائية:

- فكلما اقترب من الخط الأفقي نقصت القيمة الإحصائية للدراسة كما في C وD.
- وكلما ابتعد عنه زادت القيمة الاحصائية مثل A وB، لأن **الخط الأفقي يشير إلى عدم الفعالية الإحصائية**.

❄ كتلخيص لما سبق نجد:

- (A) فعالة إحصائياً ومنخفضة الدقة.
- (B) فعالة إحصائياً ومرتفعة الدقة.
- (C) غير فعالة إحصائياً ومنخفضة الدقة.
- (D) غير فعالة إحصائياً ومرتفعة الدقة.

هل نضع في بحثنا مسافة الثقة CI أم قيمة P؟

- يمكن للباحث أن يضع كلا القيمتين في بحثه، لكن مسافة الثقة CI تعطينا معلومات أكثر فائدة من قيمة P.
- إذا أردنا استخدام واحدة منهم فقط ببحثنا فيجب أن نستخدم مسافة الثقة.
- قيمة الـP أقل أهمية ويمكن أن نستنتجها من مسافة الثقة، لأنَّه كما قلنا فهما مكملتان لبعضهما البعض.

قياس العدد اللازم للمعالجة Number Needed To Treat (NNT):

- وجدنا منذ قليل أنَّ الخطر النسبي RR هو حاصل قسمة مقياس الخطر عند مجموعة المرضى على مقياس الخطر عند مجموعة المقارنة، وسنتعرف الآن على بعض المفاهيم الإحصائية التي تعتمد في حسابها على الخطر النسبي، وهي:
  - إنقاص الخطر النسبي Relative Risk Reduction (RRR) = 1 – الخطر النسبي RR: أي RRR = 1 – RR
  - إنقاص الخطر المطلق Absolute Risk Reduction (ARR) = الخطر في مجموعة المقارنة - الخطر في المجموعة المعالجة: أي ARR = Rc - Ra
- العدد اللازم للمعالجة NNT: يشير لعدد المرضى الذين يجب أن يتعالجوا لكي نأخذ فائدة وهل هذه الفائدة مستحقة لما قمنا ببذله من مال وجهد وغير ذلك.
- وهي أيضاً دراسة إحصائية مختلفة عن الـ P Value، ويمكن أن ندخلها بالمراجعات المنهجية بحيث تعطينا فكرة جديدة عن الدراسة الإحصائية وهي تفيد أيضاً بتقييم نتائج المعالجة.
- وبالتالي نجد أخيراً أن العدد اللازم للمعالجة:

$$NNT = \frac{1}{ARR}$$

- تم إيجاد مفهوم العدد اللازم للمعالجة حديثاً منذ حوالي 10-15 سنة.
- يهدف العدد اللازم للمعالجة إلى ما يلي:
  - تمثيل الفائدة السريرية من التجربة.
  - دراسة تكلفة الدراسة والجدوى الاقتصادية.
  - معرفة عدد المرضى الذين نحتاجهم بالدراسة لنحصل على نتائج جيدة ونمنع السلبية منها.
  - معرفة إن كانت هذه الدراسة تستحق تعميمها أم لا، وهل يمكن تعميمها أساساً.

- معرفة هل تناسب هذه الدراسة مريضي أم لا.
- هي ذات أهمية كبيرة لشركات التأمين.

❄ **مثال1:** الأولانزبين دواء للفصام وبفرض سعر علبة الدواء 15000 ل.س أي أنه عالي التكلفة، ففي حال أعطيناه لـ 100 مريض، وبعد دراسة إحصائية وجدنا أنه قد تحسن منهم 7مرضى فقط، فنكون قد أخضعنا 100 مريض للمعالجة وللتأثيرات الجانبية الكثيرة للدواء، ودفعنا الكلفة الطائلة للدواء، بالمقابل لم يتحسن سوى 7 أشخاص فهنا تبرز أهمية الـ NNT لتحديد جدوى هذه الدراسة من عدمها.

❄ **مثال2:** دواء سرطان يجب العلاج لمدة سنة لنتجنب حدوث وفاة واحدة، تم حساب العدد اللازم للعلاج فكان 20 مريضاً، والجرعة الواحدة تكلف 500$، وبالتالي المريض الواحد يُكلّف حوالي مليون دولار، إذاً هذا الدواء ليس له جدوى اقتصادية.

❄ **مثال3:** إذا كان الدواء رخيص وتأثيراته الجانبية كبيرة ونتائج معالجته قليلة، والـ RR سلبية وغيرها من المؤشرات تدل على عدم فعالية الدواء لكن هناك بعض المرضى المستفيدين من الدواء فهل نقوم باستخدام الدواء أم لا وهذا كله يُناقش في الـ NNT.

❄ **مثال4:** حاصر بيتا نستخدمه في دراسة تضم 11 مريض تحسَّن منهم واحد مع وجود تأثيرات جانبية للدواء، ففي حال أردنا تعميم استخدامه يجب أن نقرر هل يستحق أن يدفع 11 شخص أموالهم لاقتناء الدواء لكي يستفيد شخص واحد فقط من كلّ 11 وهذا ما يقرره الـ NNT.

❄ **مثال5:** لدينا 50 مريض مصابين بفصام معند، وقمنا بإعطائهم دواء لهذا المرض، لاحظنا أنَّ 20 مريض منهم فقط استجاب للدواء، وكانت تكلفة هذه الدراسة مرتفعة، فعندها علينا أن نحسب إذا كانت هذه النتيجة (استجابة 20 مريض) تستحق التطبيق وهذه الكلفة وهل يحقق هذا العدد المستفيد دراسة إحصائية تبرر لي استخدام هذا الدواء أم لا.

تطبيق النتيجة على المريض Applicability

◈ هذه الخطوة تتطلب مراجعين مقيمين على البحث وظروف وشروط إجراء الدراسة، فقبل تطبيق نتائج الدراسة على الممارسة السريرية وتعميمها على المرضى يجب التأكد مما يلي:

✪ أن تكون جمهرة الدراسة مشابهة للجمهرة التي سأقوم بالتطبيق عليها.

✪ مراعاة إمكانية تطبيق هذه الدراسة على هذا البلد، أي إمكانية التوفر في هذا البلد.

✪ مراعاة كلفة التطبيق.

✪ النظر في فوائد وأضرار هذا الإجراء.

✪ مراعاة رغبة وتفضيلات المريض في حال كان الإجراء غازي مثلاً فهل يرغب المريض بتطبيقه أم لا، وأن تكون نتائج الدراسة مماثلة لسؤال المريض PICO[25].

مثال: إجراء دراسة حول استخدام حاصرات الكلس لعلاج ارتفاع الضغط الشرياني المترافق مع قصور أو فرط نشاط درق، فبعد البحث وجدنا أنَّه لا يتوفر دراسات مطابقة لما أريد لكني وجدت دراسة بعنوان قصور أو فرط قشر كظر تجيب عن سؤالي داخلها، فهل أعتبر هذه الدراسة تماثل مريضي أم لا؟

بعد التعرف على خطوات التقييم النقدي، والوصول للمحطة الأخيرة لتعميم النتائج على المرضى والذي لا يتم إلا بدراسة المصداقية والثباتية...

المصداقية والثباتية

المصداقية Validity:

❄ أي دراسة أو استبيان يتم إجراؤه ولا يتم اختبار مصداقيته هو خاطئ وغير موثوق.

❄ تُصنّف المصداقية إلى:

1. المصداقية الخارجية External Validity (التعميم):

تتضمن كل ما تم الحديث عنه في تطبيق النتائج على المريض.

2. المصداقية الداخلية Internal Validity:

تتعلق المصداقية الداخلية بالعيّنة، وتشمل:

A. العشوائية والتعمية (التقنع).

B. المتابعة والتأثيرات الجانبية.

C. المخرجات والتحليل.

D. التحيز Bias.

[25] مثلاً مريض جاء إلى العيادة باستفسار وقمنا بوضع سؤال بحثي PICO مطابق لاستفساره وبعدها قمنا بإجراء بحث ودراسة حول هذا السؤال وفي النهاية وجدنا الجواب داخل دراسة أخرى عنوانها لا يطابق عنوان سؤالي لكنها أجابت عن السؤال.

الثباتية أو المعولية Reliability:

- تعني إمكانية إعادة نفس الدراسة بنفس الشروط مع الحصول على نفس التحليل للنتائج.
- مثال: تم إجراء استبيان من أجل كشف الاكتئاب في جامعة دمشق وحصلنا على نتائج معينة، إذا قمنا بعد مدة ما أو بذات الوقت بإعادة إجراء هذا الاستبيان بنفس الشروط يجب أن أحصل على تحليل النتائج نفسه، مع إمكانية اختلاف نسبة الانتشار، لكن الاختبار يجب أن يبقى صادقاً.

ننتقل الآن إلى التفصيل في موضوع هام يجب مراعاته في التقييم النقدي وهو التحيز...

التحيز Bias

التعريف:

- هو الاختلاف بين نتيجة الدراسة والحقيقة، وبالطبع لا نستطيع معرفة الحقيقة لكننا نحاول أن نكون قريبين منها بقدر ما نستطيع من خلال استخدام تصاميم جيدة والتنفيذ الجيد للدراسة.

أنواع التحيز:

1. التحيز المنهجي (الخطأ غير العشوائي):

- أكثر نمط مهم من التحيز.
- تكون النتيجة فيه تحت أو فوق النتيجة المتوقعة (النتيجة التي تم تقديرها).

2. التحيز غير المنهجي (الخطأ العشوائي أو خطأ المصادفة):

- يحصل بشكل متشابه في مجموعتيّ الدراسة (العلاج والشاهد).
- قابل للتوقع، وأقل أهمية من التحيز المنهجي.

أسباب التحيز:

- يوجد لدينا 10- 12 سبب للانحياز مثل الباحث أو العينة أو الدواء أو الشركة، لذلك يجب أن نتجنبهم عندما نقوم بدراسة RCTs وأن نذكر كيف قمنا بتجنبهم، كالتالي:
  - انحياز العينة: نقضي عليه بالعشوائية.
  - انحياز الأشخاص القائمين على البحث: نقضي عليه من خلال التعمية.
  - انحياز الدواء: نقضي عليه باستخدام الـ Placebo.

كما يجب أن نكون قد قمنا بكل القياسات الإحصائية مثل P Value وRR وغيرها مما درسناه في المحاضرة لنقول أنَّه لا يوجد انحياز في الدراسة.

مقياس Jadad:

- هو مقياس الانحياز الأكثر استخداماً عالمياً.
- وهو عبارة عن 3 أسئلة يجيب عنها الباحث بـ نعم أو لا، وفي حال الإجابة بنعم يأخذ نقطة "1" وفي حال الإجابة بلا لا يأخذ شيء "0" والأسئلة هي كالتالي:
  1. هل كانت الدراسة عشوائية؟
  2. هل تضمنت الدراسة تعمية ثنائية؟
  3. هل كان هنالك تحيز انسحاب[26]؟
- بعد ذلك يوجد سؤالين يجب الإجابة عنهما، ويتم إعطاء نقطة "1+" في حال تمت الإجابة بنعم على السؤال الأول، ويتم حذف نقطة "1-" في حال تمت الإجابة بنعم على السؤال الثاني، وهما:
  1. هل كانت العشوائية والتعمية مناسبة للدراسة؟
  2. هل كانت العشوائية والتعمية غير مناسبة للدراسة؟
- وبالنتيجة نقوم بجمع النقاط ويكون الناتج ضمن المجال (0-5).
- في حال كانت النتيجة أقل أو تساوي 2 تكون الدراسة ضعيفة.

توضيح خارجي:

- في البند الثالث من هذه الفقرة ذكرنا أنه يتم حذف نقطة "1-" في حال تمت الإجابة بنعم، لكن هنا يتم حذف النقطة بشرط ألا يقل مجموع النقط عن "0" بعد هذا الحذف أي لا يجوز أن تكون النتيجة سالبة، وأقل نتيجة بالمجمل هي "0".

أنماط التحيز في دراسات الـ RCTs:

- يوجد العديد من الأنماط للتحيز في دراسات الـ RCTs وذلك **تبعاً لطور الدراسة** كالتالي:

1. خلال طور التخطيط للدراسة:

- **تحيز باختيار السؤال البحثي** Choice Of Question Bias.
- **تحيز بالتنظيم** Regulation Bias.

[26] أي انسحاب بعض الأشخاص من الدراسة، مما يؤثر على مسارها ويؤدي للانحياز.

- تحيز باختيار تصميم خاطئ Wrong Design Bias.

2. خلال القيام بالدراسة:

- تحيز بالانتقاء Selection Bias.
- تحيز بالملاحظة Observation Bias.
- تحيز باختيار الجمهرة Population Choice Bias.
- تحيز بخيار التداخل Intervention Choice Bias.
- تحيز بالمجموعة الشاهدة Control Group Bias.
- تحيز بالنتائج Outcome Bias.

3. خلال إظهار تقرير الدراسة (النتائج):

- تحيز الانسحاب Withdrawal Bias.
- تحيز بكتابة تقارير انتقائية Selective Reporting Bias.
- تحيز بالتزوير والنصب Fraud Bias.

بعد التعرف على مفهوم التحيز، ننتقل إلى عرض تقييم الدراسات على شكل استبيان من 12 سؤال لتنظيم الأفكار لدى الناقد...

تقييم دراسات الـ RCTs

❖ تم تلخيص كل ما سبق على شكل استبيان مقسم لثلاثة أقسام من 12 سؤال لتقييم الدراسة:

هل كانت النتيجة ذات قيمة؟

☒ **في بداية الدراسة:**

1. هل المرضى عشوائيين؟
2. هل العشوائية مخفية؟
3. هل استخدمت نفس عوامل التشخيص في كلتا المجموعتين؟

☒ **خلال الدراسة:**

4. هل كانت الدراسة معماة وإلى ماذا امتدت؟

☒ **في نهاية الدراسة:**

5. هل كانت المتابعة كاملة؟

6. هل تم تطبيق الـ ITT؟
7. هل تم إيقاف الدراسة مبكراً؟

ماذا كانت النتائج؟

8. كم كان مدى تأثيرات العلاج؟
9. ماهي دقة التقدير لتأثيرات العلاج أي فاصلة الثقة (CI)؟

كيف أستطيع تطبيق النتائج لرعاية مريضي؟

10. هل كان المرضى بالدراسة مشابهين لمرضاي؟
11. هل تم وضع جميع المرضى والنتائج المهمة بعين الاعتبار؟
12. هل الفوائد المحتملة للعلاج كانت تستحق الضرر والكلفة التي تم وضعها؟

يوفر موقع كوكرن صفحة استبيان يقوم الباحث بملئها ويقدمها للحكام على دراسته لكي يقيموا من خلالها هل انحاز في دراسته أم لا.

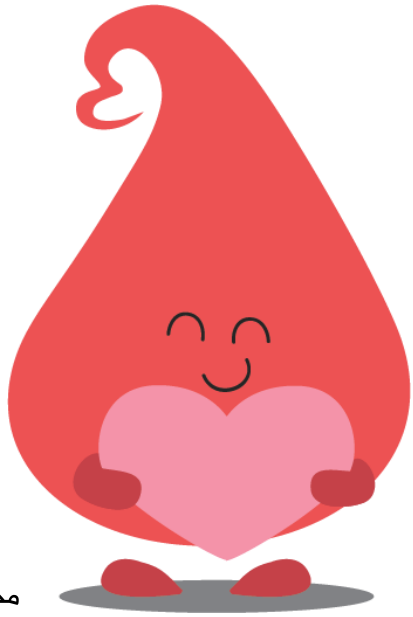
إلى هنا أصدقائي نكون وصلنا إلى نهاية رحلتنا مع هذه المادة المُمتعة التي تعلمنا فيها كيف نكون على بيّنة بكلّ قرار طبي نتّخذه كأطباء لنكون على قدرٍ من الإخلاص والمسؤولية أمام الله أولاً والمريض ثانياً

نأمل أن نكون قد وُفقنا في إيصال المعلومة بالشكل الأنسب

فريق الطب المسند يتمنّى لكم امتحاناً موفّقاً

لا تنسونا من صالح الدعاء

والسلام عليكم ورحمة الله وبركاته*-*



RBCs