



Modul 9

FEB 326-Evidence-Based Practice Fisioterapi

Materi 9

Penilaian Kritis Pada Studi Terkait Penentuan Diagnosis

Disusun Oleh

Wahyuddin

UNIVERSITAS ESA UNGGUL

2019

A. Pendahuluan

Penegakan diagnosis adalah komponen penting dalam aplikasi *evidence-based practice* fisioterapi. Untuk dapat menghasilkan diagnosis yang akurat, diperlukan pemahaman terkait jenis studi dan melakukan kritisi pada studi tersebut.

B. Kompetensi Dasar

Kompetensi dasar yang harus dimiliki oleh mahasiswa dan mahasiswi pada materi ini adalah pemahaman tentang proses kritisi studi terkait penegakan diagnosis

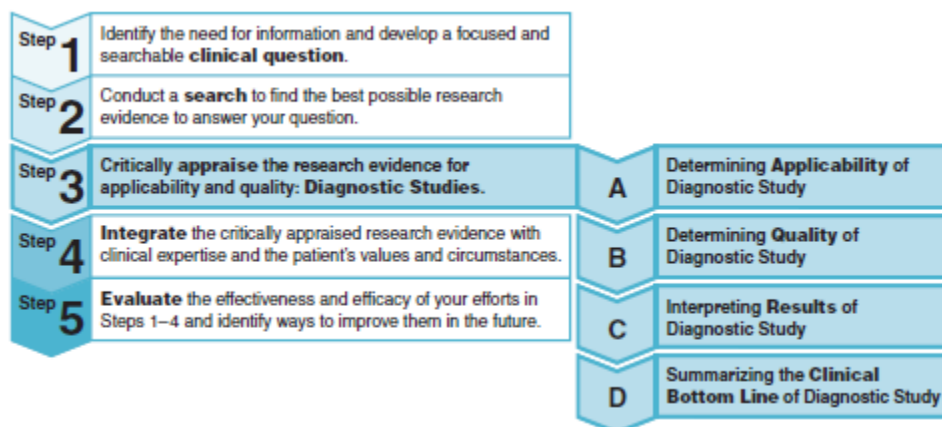
C. Kemampuan Akhir Yang Diharapkan

Setelah mempelajari materi ini, diharapkan semua mahasiswa dan mahasiswi mampu mengkritisi studi terkait penentuan diagnosis dan memahami pentingnya hal tersebut dalam aspek praktik klinis

D. Uraian Materi

Proses Diagnostik dalam Fisioterapi

Proses diagnostik fisioterapi meliputi riwayat pasien, review sistem serta diinformasikan melalui penggunaan tes dan pengukuran. Bagian ini berfokus pada proses penilaian literatur diagnostik tes dan pengukuran dan penerapan klinis hasil penelitian diagnostik (gambar 1).



Gambar 1. Penilaian Kritis Studi Penelitian Diagnostik

Pemahaman terkait bukti-bukti mengenai tes diagnostik dan mengintegrasikan hasil tes diagnostik, pengalaman klinis, dan tujuan pasien sangat penting untuk untuk mendapatkan hasil praktik terbaik. Komunikasi hasil proses diagnostik terhadap pasien dan profesional lainnya memerlukan pengetahuan mendalam tentang tes diagnostik yang valid dan reliabel serta interpretasi yang tepat dari tes tersebut. Fisioterapis melakukan pemeriksaan pada pasien dan mengevaluasi hasil pemeriksaan, menghasilkan diagnosis, rencana tindakan, dan kemungkinan prognosis. Proses diagnostik dalam terapi fisik telah dibahas di berbagai literatur. Ketika terlibat dalam proses diagnostik, kita menentukan intervensi fisioterapi yang sesuai dan jika diperlukan kebutuhan untuk rujukan ke profesional lain. Akses langsung ke fisioterapi berdampak kepada peningkatan tanggung jawab profesional. Bagian dari tanggung jawab

tersebut adalah diagnosis yang akurat dari kondisi dalam lingkup praktek fisioterapi dan rujukan ke profesional lain secara tepat

Sebelum akses langsung, fisioterapis fisik mungkin melakukan tindakan berdasarkan resep atau rujukan dari dokter. Rujukan ini berfungsi sebagai sumber rujukan utama bagi fisioterapis, tetapi sebagian besar rujukan ini tidak memiliki diagnosis spesifik. Misalnya, mungkin ditulis sebagai "nyeri siku," "peningkatan kapasitas aerobik", atau "paska stroke," tanpa diagnosis mengenai etiologi sebagai penyebab masalah atau hubungannya dengan gerakan. Fisioterapis bertanggung jawab untuk mendiagnosis dengan tepat masalah gerakan dan menentukan intervensi yang tepat, kemungkinan prognosis, dan, jika perlu, rujukan ke profesional lainnya.

Prinsip dasar dari setiap tes diagnostik yang digunakan dalam fisioterapi adalah bahwa hal itu akan dapat membedakan antara orang yang memiliki masalah terkait gerak dan juga fungsi dengan orang-orang yang tidak memiliki masalah tersebut. Meskipun ini terdengar jelas, namun persoalannya kompleks. Jika kita memiliki sakit tenggorokan, seorang dokter mungkin akan menggunakan rapid antigen throat culture bakteri streptokokus untuk mendiagnosis infeksi streptokokus. Kemungkinan besar akan mengasumsikan bahwa hasil tes ini menentukan bahwa kita memiliki radang tenggorokan atau tidak. Namun, setiap tes diagnostik memiliki serangkaian nilai-nilai yang positif dan berbagai nilai-nilai yang negatif. Tes cepat ini untuk radang tenggorokan mendeteksi 80% sampai 90% dari infeksi streptokokus, dengan demikian hilang 10% sampai 20% dari kasus yang ada. Tes ini mungkin tidak melihat adanya infeksi streptokokus bahkan jika itu adalah hadir (false-negative). Namun, beberapa tes positif juga mungkin tidak akurat (false-positives).

Tidak ada tes diagnostik yang 100% akurat dalam ruling in atau ruling out suatu diagnosis. Pelayan medis akan menggunakan rapid antigen throat culture berdasarkan riwayat dan gejala klinis untuk menentukan apakah pasien memerlukan perawatan untuk infeksi streptokokus dan kemungkinan jika pengujian tambahan diperlukan. Setiap pertanyaan terkait riwayat, pengamatan selama review sistem, dan setiap pilihan pengujian atau mengukur harus didasari oleh hipotesis awal tentang masalah. Oleh karena itu, informasi yang kita cari harus diarahkan dari hal umum ke spesifik. Kita menentukan tes dan pengukuran apa yang akan paling bermanfaat ketika jenis informasi yang kita butuhkan berdasarkan sumber yang valid.

Sebagai contoh, pertimbangkan seorang pasien dengan diagnosis spesifik nyeri bahu. Pasien melaporkan riwayat dan gejala yang konsisten dengan misalnya labral tear, supraspinatus tear, atau supraspinatus tendonitis. Hipotesis mengenai masalah gerak didasari oleh laporan pasien dan kemudian diuji dengan pemeriksaan. Tes anterior slide yang positif dapat mendukung hipotesis labral tear. Hasil setiap tes yang dipilih dalam pemeriksaan harus secara langsung relevan baik ruling in or to ruling out terhadap diagnosis tertentu. Seseorang yang mendiagnosis secara efisien akan bertanya, "Sejauh mana hasil dari tes ini akan membantu keputusan saya?" Untuk menentukan jawaban atas pertanyaan ini kita perlu mengetahui uji klinis yang valid dan reliabel. Dalam contoh kasus labral tear, kita harus mempertimbangkan pertanyaan, "Apa akurasi anterior slide test terkait ruling in atau ruling out labral tear?". Jawaban pertanyaan ini didapatkan dari hasil studi tersebut.

Pertanyaan-Pertanyaan Terkait Diagnosis di Klinik

Diagnosis dalam fisioterapi merupakan hasil dari integrasi proses pemeriksaan dan evaluasi. Diagnosis yang tepat merupakan dasar untuk memberikan pilihan intervensi dan potensi prognosis terhadap pasien. Namun, banyak dari tes yang digunakan dalam fisioterapi mungkin tidak telah sepenuhnya dievaluasi kualitas dan reliabilitasnya. Kualitas suatu tes dilakukan untuk tujuan tertentu dan sering dikaitkan dengan pertimbangan masalah pasien tertentu.

Untuk setiap uji diagnostik yang digunakan dalam pemeriksaan, kita harus mengetahui jawaban atas pertanyaan-pertanyaan berikut:

- Apa informasi diagnostik yang saya cari?
- Apakah suatu tes akan membantu saya dalam mengkonfirmasi atau menyangkal hipotesis klinis mengenai masalah terkait gerak pada pasien saya?
- Bagaimana hasil tes ini akan mempengaruhi rekomendasi terhadap treatment saya?
- Apa yang saya ketahui tentang karakteristik-karakteristik tes yang saya gunakan?
- Apa yang harus saya ketahui tentang karakteristik-karakteristik tes yang saya gunakan?

Pencarian Literatur: Desain Studi Spesifik pada Tes Diagnostik

Cohort design adalah desain studi paling umum untuk pengembangan tes diagnostik. Prinsip yang sama untuk pencarian literatur yang dijelaskan dalam materi-materi sebelumnya juga berlaku untuk pencarian literatur terkait diagnostik. Namun, jenis desain yang digunakan dalam studi tes diagnostik dan langkah-langkah cohort designs, bukan randomized clinical trials (RCTs). Prospective cohort designs lebih disarankan dibanding retrospective designs untuk evaluasi diagnostik tests. Meskipun dimungkinkan untuk menggunakan retrospective designs, desain ini mempunyai tantangan besar dalam hal keabsahan informasi.

Contoh hipotetis retrospective design akan membantu menjelaskan kebutuhan untuk prospective designs. Misalnya bahwa peneliti ingin mengetahui jika hasil Lachman test yang biasanya digunakan di klinik untuk pasien dengan dugaan anterior cruciate ligament (ACL) tear menghasilkan hasil yang sama seperti magnetic resonance images (MRIs) pada pasien tersebut. Mereka mengkonsulkan berdasarkan catatan medis pasien untuk menentukan hasil kedua tes tersebut. Namun, tidak semua pasien dilakukan Lachman tes, atau mungkin tes tidak dicatat, atau ini mungkin tidak jelas dalam catatan. Selain itu, reliabilitas antara dokter yang melakukan Lachman test atau MRI yang rendah, sehingga mustahil untuk menentukan jika tes yang diberikan dengan valid dan atau reliabel. Prospective studies tes diagnostik dilakukan untuk pengumpulan data pada uji klinis (Lachman test) dan perbandingan dengan tes yang bersifat gold standard (MRI).

Pengembangan tes diagnostik yang valid dan reliabel memerlukan penggunaan dua tes, yaitu tes indeks dan gold standard pada setiap subjek. Dalam sebuah studi penelitian, prosedur pengumpulan data dilakukan dengan studi desain yang valid, misalnya terkait reliabilitas penguji untuk kedua tes tersebut. Peserta studi setuju untuk kedua prosedur sebagai bagian dari penelitian. Hal penting adalah perencanaan ke depan yang melekat dalam

prospective study. Subjek dapat dipilih dalam satu titik waktu (prospective and cross-sectional) dan data diperoleh sekali dan dalam jangka waktu singkat atau sampel pada satu waktu titik dan diikuti selama jangka waktu lebih lama untuk menentukan jika kondisi target terkonfirmasi (prospective and longitudinal).

Proses Penilaian Studi

Seperti yang dibahas sebelumnya, penilaian adalah langkah ketiga dalam lima langkah proses evidence based practice (EBP). Dalam pembahasan ini kita dapat mengembangkan keahlian untuk menilai literatur diagnosis dan menyelesaikan empat bagian yang sangat penting untuk penilaian literatur diagnostik yaitu:

- menentukan kemampuan penerapan studi;
- menentukan kualitas studi;
- menginterpretasikan hasil studi

Ada berbagai pendekatan untuk proses penilaian untuk studi diagnostik. Salah satu alat penilaian yang saat ini digunakan dalam literatur fisioterapi adalah Quality Assessment of Diagnostic Accuracy Studies (QUADAS). QUADAS meliputi 14 pertanyaan untuk acuan penilaian penerapan dan kualitas studi diagnostik. Setiap pertanyaan dicatat sebagai "Ya," "tidak," atau "tidak jelas." Jumlah total jawaban "ya" dapat diringkas untuk skor kualitas pada setiap studi diagnostik yang dinilai.

Walaupun proses pemberian skor kualitas untuk penggunaan diagnostik studi systematic reviews tidak dianjurkan, beberapa penulis menyarankan skor total dapat digunakan untuk mendapatkan estimasi kualitas studi diagnostik. Penilaian lain yang dalam literatur-literatur fisioterapi adalah Standards for Reporting of Diagnostic Accuracy (STARD). Penggunaan penilaian tersebut, digunakan untuk secara sistematis pada laporan penelitian studi diagnostik untuk publikasi dan tidak langsung berlaku untuk proses penilaian untuk masing-masing studi diagnostik.

Bagian A: Menentukan Kemampuan Aplikasi Studi Diagnostik

Pertanyaan 1: Apakah seluruh spektrum pasien terwakili dalam sampel studi? Apakah pasien saya merepresentasikan pasien-pasien lain?

Spektrum dampak penyakit, kecacatan atau cedera harus terwakili dalam studi tes diagnostik. Apabila seluruh spektrum tidak diwakili, kemudian penerapan ujian akan terbatas dan ini harus diakui oleh para penulis. Studi harus menyertakan populasi klinis yang relevan untuk tes, tetapi juga berkaitan dengan pasien. Pertanyaan ini kemudian penting untuk penerapan pada pasien dan kualitas penelitian.

Sebagai contoh, Berg scale telah divalidasi pada sampel dari orang yang memiliki gangguan gerak sebagai akibat dari stroke dan juga pada mereka yang tanpa gangguan gerak dan tanpa stroke. Skala yang sama belum divalidasi dengan orang-orang dengan penyakit Parkinson. Meskipun orang-orang dengan penyakit Parkinson memiliki gangguan gerak dan keseimbangan, hal itu belum ditentukan bagaimana mereka menggunakan skala tersebut. Artinya jika terdapat situasi seperti itu berarti tes tersebut mungkin tidak tepat diaplikasikan pada mereka yang mengalami penyakit Parkinson. Dan keterbatasan tersebut harus diakui ketika menginterpretasikan hasil tes. Suatu tes tertentu harus digunakan untuk tujuan yang telah ditetapkan.

Pengukuran outcome yang digunakan dalam penelitian intervensi harus valid mengukur hasil outcome yang diharapkan. Misalnya, peningkatan aktifitas kehidupan sehari-hari adalah tujuan dari intervensi, kemudian menggunakan ukuran kekuatan tidak akan memberikan wawasan tentang dampak yang diharapkan dari intervensi. Tes diagnostik harus sama dalam hal evaluasi dan penerapan. Sebagai contoh, norm reference developmental motor scale tepat digunakan untuk menentukan kelayakan program intervensi awal. Namun, jenis tes ini tidak berguna dalam menentukan klasifikasi anak intervensi tertentu. Misalnya, seorang anak dengan Down syndrome dan seorang anak dengan cerebral palsy mungkin keduanya skor standar deviasi di bawah rata-rata Bayley Scales of Infant Development-II dan kriteria ini dapat membuat setiap anak mempunyai eligibilitas untuk layanan intervensi dini. Namun, skor ini tidak akan membantu fisioterapis dalam penentuan rekomendasi khusus untuk intervensi untuk setiap anak, atau akan itu mengkonfirmasi diagnosis down syndrome atau cerebral palsy.

Bagian B: Penentuan Kualitas Studi Diagnostik

Idealnya, tes dan pengukuran yang digunakan dalam fisioterapi harus divalidasi melalui penelitian klinis. Hal ini berdampak pada pada semakin meningkatnya literatur terkait. Ada beberapa hal penting dari penelitian tes diagnostik yang harus dipertimbangkan pada proses penilaian sebelum penerapan hasil studi tersebut. Seperti halnya dengan intervensi, atau jenis lain dari penelitian, kualitas penelitian harus dinilai sebelum tes diagnostik apakah dapat diterapkan di klinik. Ada aspek umum studi kualitas untuk studi diagnostik, prognostik dan intervensi. Namun, ada aspek validitas yang unik untuk studi diagnostik.

Pertanyaan 2: Apakah ada perbandingan secara independen dan blinding dengan diagnosis yang bersifat gold standard?

Seperti yang telah dijelaskan sebelumnya, kualitas (kadang-kadang disebut sebagai validitas) tes atau pengukuran didapatkan melalui perbandingan hasil tes diagnostik (disebut sebagai indeks test) terhadap hasil tes lain, yang disebut sebagai gold standard. Indeks dan gold standard test harus memiliki tujuan yang sama. Sebuah gold standard adalah tes terbaik yang diterima secara luas sebagai yang tersedia. Banyak tes yang dikembangkan dalam fisioterapi yang digunakan untuk diagnosis muskuloskeletal telah dibandingkan dan divalidasi terhadap tes pencitraan, misalnya, MRI. Tes valid yang dilakukan oleh fisioterapis mungkin menghemat waktu dan biaya jika dapat menghasilkan informasi yang relatif sama dibanding menggunakan tes-tes berbiaya tinggi yang tidak selalu tersedia

Pertanyaan 3: Apakah tes diagnostik yang dilakukan oleh satu atau lebih penguji reliabel terhadap hasil pengujian dengan gold standard?

Penelitian mengenai validitas pengujian harus mencakup blinding aplikasi tes yang dilakukan dengan gold standard. Hal ini memerlukan setidaknya dua penguji, dimana setiap menyelesaikan hanya salah satu tes. Setiap tes harus diselesaikan tanpa mengetahui hasil tes lain dan idealnya tanpa pengetahuan tentang tujuan studi. Salah satu ciri penting tes berguna adalah bahwa hal itu bersifat reliabel. Reliabilitas pengujian dipengaruhi oleh stabilitas fenomena sedang diuji (misalnya, keseimbangan, kualitas hidup) maupun oleh

kemampuan penilai untuk mengulang tes tanpa adanya perubahan yang signifikan.

Telah terdapat banyak penelitian reliabilitas pada uji klinis kami. Namun, sebagian besar uji reliabilitas dilakukan dengan sampel yang tidak memiliki masalah gerak. Aspek reliabilitas pada orang sehat biasanya bergerak ini tidak memadai untuk studi tes diagnostik yang dirancang untuk populasi pasien. Kedua intra-rater dan inter-rater harus ditetapkan terhadap tes yang dilakukan dan gold standard sebagai pembanding. Keandalan harus ditetapkan pada tipe pasien yang sesuai dan rentang kesulitan tes termasuk dalam studi tersebut.

Pertanyaan 4: Apakah semua subjek menerima kedua tes (indeks dan gold standard)?

Kedua tes harus dilakukan untuk melihat outcome masing-masing. Seperti disebutkan di atas, dua atau lebih profesional secara blinding melakukan perbandingan ini dan kemudian hasil uji dianalisis. Misalnya, fisioterapis akan menyelesaikan Lachman tes pada peserta dalam studi penelitian klinis, dan hasilnya akan dibandingkan dengan subjek yang sama hasil dari MRI (gold standard) yang diinterpretasikan oleh ahli radiologi.

Idealnya, kedua tes akan menghasilkan informasi diagnostik yang sama. Ini berarti, misalnya bahwa meskipun seseorang memiliki hasil Lachman test negatif yang mengindikasikan ACL tidak mengalami masalah, orang tersebut harus melakukan MRI untuk mengkonfirmasi temuan tes negatif pada Lachman test.

Bagian C: Interpretasi Hasil Studi Diagnostik

Pertanyaan 5: Apakah tes diagnostik diinterpretasikan secara independen pada hal terkait informasi klinis?

Fisioterapis menggunakan tes dan pengukuran dalam hubungannya dengan riwayat pasien dan pengalaman klinis. Ketika tes dikembangkan informasi tambahan ini tidak selalu tersedia untuk orang-orang yang melakukan tes indeks dan gold standard test. Fisioterapis ingin mengetahui apakah tes yang dilakukan pada suatu studi akan berkontribusi terhadap proses diagnostik selain informasi riwayat dan keadaan pasien.

Informasi ini ditentukan dari penelitian diagnostik yang dikendalikan secara hati-hati. Misalnya, jika tes klinis keseimbangan dibandingkan dengan tes gold standard, maka setiap tes banyak diberikan tanpa pengetahuan tentang hasil tes lain dan tanpa pengetahuan tentang karakteristik klinis tambahan subyek.

Pertanyaan 6: Apakah secara klinis menggunakan statistik untuk analisis dan interpretasi pada aplikasi klinis?

Membandingkan suatu tes terhadap gold standard meliputi perhitungan statistik diagnostik. Ada banyak jenis statistik yang berbeda yang dapat digunakan untuk mencerminkan akurasi dan kegunaan tes untuk diagnosis fisioterapi. Beberapa statistik yang termasuk dalam bagian ini juga digunakan dalam penelitian prognostik. Statistik telah dimasukkan yang dapat diterapkan untuk proses diagnostik klinis.

Statistik yang Umum Digunakan pada Studi Diagnostik

Pembahasan pada bagian ini fokus pada aspek statistik pada studi diagnostik yang meliputi:

- Sensitivity
- Specificity
- Receiver operating curves
- Positive and negative predictive values
- Positive and negative likelihood ratios

Perhitungan hal-hak di atas dapat dilakukan melalui tabel 2 × 2 pada tabel 1 berikut ini:

	GOLD STANDARD POSITIVE	GOLD STANDARD NEGATIVE
Clinical test Positive	A True positive	B False-positive
Clinical test Negative	C False-negative	D True negative

Table cell A = Patients with positive Lachman and positive MRI tests are true positives

Table cell B = Patients with positive Lachman and negative MRI tests are false-positives

Table cell C = Patients with negative Lachman and positive MRI tests are false-negatives

Table cell D = Patients with negative Lachman and negative MRI tests are true negatives

Tabel 1. Tabel 2x2

Memahami tabel 2 × 2 tables memberikan pemahaman yang lebih jelas tentang sensitivity, specificity, predictive values, dan likelihood ratios. Tabel 1 meliputi hasil tes klinis yang diinvestigasi (positif atau negatif) dan hasil partisipan yang sama berdasarkan gold standard test. Gold standard secara efektif digunakan untuk menentukan “kebenaran” tentang ada atau tidaknya kondisi. Karena itu, “kebenaran” adalah sebaik gold standard yang digunakan sebagai pembanding.

Sebagai contoh, suatu studi meliputi 60 sampel dengan riwayat trauma dan cocok dengan gambaran ACL tear. Semua pasien mempunyai kedua hasil berupa Lachman test yang dilakukan oleh fisioterapis dan MRI lutut yang dievaluasi oleh radiologis. Kedua hasil tersebut dimasukkan ke dalam tabel (tabel 1) dan kita dapat menentukan kesepakatan atau ketidaksepakatan antara dua tes. Seperti yang terlihat pada tabel 1, perhitungan statistik untuk diagnostik tersebut menghasilkan suatu dichotomous variables apakah positif atau negatif. Di sini ditentukan bahwa suatu tes mempunyai specific cut off point baik dengan hasil positif atau negatif

Lachman test akan menghasilkan baik positif atau negatif, dan MRI akan menghasilkan kesimpulan apakah terdapat ACL tear atau tidak. Ahli radiologi perlu menentukan kriteria untuk mendefinisikan hasil MRI positif versus MRI negatif. Setiap pasien data akan dibuatkan kode ke dalam salah satu kotak dari empat dalam tabel. Kolom dalam tabel 2 × 2 adalah hasil gold standard test,

dalam contoh ini MRI. Baris dalam tabel 2 × 2 adalah hasil uji klinis yang diteliti. Idealnya, semua pasien harus true positives (tabel 1 A) atau true negatives (tabel 1 D). True positive menunjukkan bahwa orang pasti memiliki kondisi, sedangkan true negative menunjukkan bahwa orang pasti tidak memiliki kondisi. Hal ini jarang terjadi bahkan dengan tes diagnostik terbaik. False positive (tabel 1 B) menunjukkan bahwa orang tidak memiliki kondisi seperti diuji dengan gold standard, tapi positif pada tes uji klinis. False-negative (Lihat tabel 5.1 C) menunjukkan sebaliknya, seseorang dengan hasil positif pada gold standard, tetapi negatif pada uji klinis. Statistik diagnostik sensitivity dan specificity harus dipertimbangkan untuk mengestimasi kualitas tes. Secara kuantitatif harus terdapat kesesuaian antara tes yang diteliti dengan gold standard.

Sensitivity

Sensitivity menunjukkan akurasi uji klinis secara tepat untuk mengidentifikasi masalah yang ditentukan dengan gold standard. Menggunakan contoh Lachman test dan MRI untuk diagnosis ACL rupture, setiap subjek ditetapkan ke kotak yang sesuai tergantung pada kombinasi hasil tes tersebut (tabel 2).

Sensitivity tes adalah jumlah total pasien yang memiliki Lachman positif dan tes MRI positif (true positive) dibagi dengan total jumlah pasien dengan hasil MRI positif. Berdasarkan formula, maka rumus sensitivity adalah:

$$[A / (A + C)]$$

	MRI POSITIVE	MRI NEGATIVE
Lachman test Positive	A True positives	B False-positives
Lachman test Negative	C False-negatives	D True negatives

A = Patients with positive Lachman and positive MRI tests are true positives
 B = Patients with positive Lachman and negative MRI tests are false-positives
 C = Patients with negative Lachman and positive MRI tests are false-negatives
 D = Patients with negative Lachman and negative MRI tests are true negatives

Tabel 2. Kombinasi Hasil Tes

Pertanyaan selanjutnya adalah dari semua pasien dengan hasil MRI positif, berapa banyak (persentase) memiliki Lachman test positif? Ekspresi alternatif adalah sebagai berikut: seberapa sensitif Lachman test dapat menentukan ACL rupture seperti yang ditentukan melalui hasil pemeriksaan MRI?

Dengan asumsi MRI menghasilkan diagnosis yang akurat apakah terdapat ACL rupture atau tidak, kita menentukan sensitivity Lachman test dalam menentukan ACL rupture. Jika Lachman dan MRI keduanya mendiagnosis masalah yang sama secara tepat, maka kita hanya perlu satu atau tes lain, tetapi tidak keduanya. Kita yakin bisa menggunakan tes yang lebih murah dan tidak memakan banyak waktu seperti Lachman test untuk menentukan integritas ACL

dan meminimalkan kebutuhan MRI untuk penegakan diagnosis. Namun, sebelum kita dapat menentukan pilihan tes tersebut, kita harus mengetahui specificity tes ini, untuk mengidentifikasi semua hasil secara tepat baik negatif serta positif.

	MRI POSITIVE	MRI NEGATIVE
Lachman test Positive	A True positives n = 39	B False-positives n = 10
Lachman test Negative	C False-negatives n = 3	D True negatives n = 8
	<i>Sensitivity</i> A + A/C 39/39 + 3 = 0.93	<i>Specificity</i> D/D + B 8/8 + 10 = 0.44

Tabel 3. Perhitungan Sensitivity dan Specificity

Menghitung sensitivity secara matematika pada tabel 3 di atas, didapatkan hasil sebagai berikut:

$$\begin{aligned} & A/(A + C) \\ & 39/(39 + 3) \\ & 39/42 = 0.93 \end{aligned}$$

Apa jenis pernyataan klinis kita tentang nilai sensitivitas tersebut? Misalnya: "Lachman test positif secara tepat mengidentifikasi 93% pasien dengan ACL torn." Kita kemudian dapat menyimpulkan dari statistik ini bahwa Lachman test secara tidak tepat mengidentifikasi 7% pasien mengalami torn ACL ketika ada tear yang didiagnosis dengan menggunakan MRI.

Specificity

Specificity menunjukkan kemampuan suatu tes mengidentifikasi tidak adanya masalah. Hal itu diperoleh dari jumlah pasien dengan Lachman test yang negatif dan MRI negatif atau true negatives (tabel 3 D) dibagi dengan jumlah total pasien dengan hasil MRI negatif.

$$D/(B + D)$$

Ini membantu kita menjawab pertanyaan berikut: "dari semua pasien dengan MRI negatif, berapa banyak (persentase) memiliki Lachman test negatif?" Ekspresi alternatif adalah sebagai berikut: "seberapa spesifik Lachman test dalam mengidentifikasi nonruptured ACL seperti yang ditentukan oleh MRI?"

Dalam klinik, keuntungan utama dari dengan tes spesifik yang baik adalah bahwa pasien tidak akan membutuhkan referral lebih lanjut untuk MRI untuk diagnosis jika Lachman test negatif (menggunakan contoh di atas). Perhitungan rumus specificity adalah sebagai berikut:

$$\begin{aligned} & D/(D + B) \\ & 8/(8 + 10) \\ & 8/18 = 0.44 \end{aligned}$$

Apa pernyataan yang dapat dibuat tentang nilai specificity ini? Misalnya: "Lachman test yang negatif mengidentifikasi secara tepat 44% pasien tanpa torn ACL." Kita kemudian dapat menyimpulkan dari statistik ini hanya bahwa tes Lachman negatif secara tidak tepat mengidentifikasi 56% pasien tidak mempunyai torn ACL robek ketika terdapat yang diidentifikasi berdasarkan pada gold standard (MRI).

Sensitivity dan specificity adalah statistik diagnostik yang dapat membantu kita dalam memilih uji klinis yang bermanfaat dalam praktik klinis. Kedua statistik ini tidak langsung tergantung pada frekuensi relatif masalah dalam populasi, tetapi pada spektrum masalah yang diwakili oleh sampel yang digunakan dalam penelitian. Sampel yang digunakan untuk menentukan sensitivity dan specificity sangat penting dalam memperoleh nilai dan dalam menentukan manfaat tes untuk pasien-pasien tertentu. Cleland membuat daftar nilai-nilai uji rentang lebar Lachman dilaporkan dalam literatur untuk sensitivity (0,65-0,99) dan specificity (1,0-0,42). Review dari dua studi yang dikutip oleh Cleland menyelidiki tes Lachman utilitas dalam mendeteksi ACL rupture menggambarkan pentingnya sampel penelitian, metode dan aplikasi ketika mempertimbangkan nilai sensitivity dan specificity dari suatu tes.

Dalam sebuah studi oleh Katz dan Fingerth, 23 pasien direkrut ketika mereka diarahkan untuk arthroscopy. Lachman positif dalam studi ini padatitik akhir untuk tibial translation atau positif subluxation. Mereka melaporkan sensitivity (0,82) dan specificity tinggi (0,97). Sebagai perbandingan, subyek dalam sebuah studi oleh Cooperman et al²⁴ direkrut ketika mereka ke fisioterapi untuk evaluasi awal lutut. Lachman test telah dinilai pada skala empat poin dengan penambahan dipisahkan oleh 5 mm atau kurang. Sensitivity (0,65) dan specificity (0,42) yang jauh lebih rendah daripada yang dilaporkan dalam studi Katz dan Fingerth.

Lachman mungkin memiliki nilai-nilai yang lebih tinggi (deteksi lebih baik) ketika mendeteksi cedera ACL yang memiliki probabilitas tinggi sebelum tes ACL rupture. Misalnya ketika subjek sudah dirujuk untuk arthroscopy dibandingkan dengan subjek yang dirujuk ke fisioterapi untuk setiap masalah lutut. Tes juga mungkin lebih sensitif dan spesifik ketika tanda-tanda esktrim Lachman test digunakan untuk menentukan temuan positif.

Receiver Operating Characteristic (ROC) Curve

Satu keputusan penting dalam menghitung sensitivity dan specificity adalah pilihan cutoff score yang digunakan untuk menentukan nilai tes positif dan negatif. Untuk mengkonstruksi dan memvalidasi tes diagnostik, cut point ditentukan yang memberikan hasil tes dichotomous (positif atau negatif). Secara klinis, cut point terbaik ditentukan dengan mempelajari efek dari berbagai cut point data pada sensitivity dan specificity. Receiver operating characteristic (ROC) curve adalah representasi grafis dari nilai sensitivity dan specificity yang dihasilkan dari serangkaian cut points nilai ini didasarkan pada premis bahwa false-positive (pasien tes positif, tetapi tidak memiliki kondisi) sama pentingnya dengan false-negative (pasien tes negatif, tapi sebenarnya memiliki kondisi).

Contoh dari tes medis mungkin bisa memperjelas pentingnya keseimbangan antara sensitivity dan specificity dari suatu tes. Jika tes streptokokus tenggorokan negatif, tetapi pasien sebenarnya memiliki radang tenggorokan (false negative), antibiotik mungkin tidak diresepkan oleh dokter.

Meskipun waktu pemulihan mungkin lebih lambat, orang yang sehat dapat sembuh dari infeksi radang tenggorokan, tetapi ada risiko, meskipun jarang, infeksi streptokokus terdeteksi menuju rematik fever. False-negative pada orang tua dengan banyak co-morbidities dapat mengancam kehidupan. Dalam situasi ini, sensitivitas tes ini penting untuk menghindari konsekuensi serius. Jika seorang pasien tes positif untuk radang, tetapi tidak memiliki infeksi streptokokus (false positive) kemudian antibiotik mungkin diresepkan ketika mereka tidak perlu.

Contoh lain dari tindakan fisioterapi menyoroti kebutuhan untuk menanyakan, "seberapa penting hasil pengujian?" Jika seorang pasien tes negatif untuk risiko jatuh, tetapi benar-benar mempunyai risiko untuk jatuh (false-negative), tindak lanjut penting mungkin tidak terjadi. Ini bisa menjadi situasi yang berpotensi mengancam. False positive dalam mendeteksi risiko jatuh mungkin tidak menjadi semacam alarm pada pasien, tapi pengujian tambahan menggunakan tes lebih spesifik dapat menjadi rule out masalah. Dengan pengujian untuk risiko jatuh, sensitivitas tes mungkin ditekankan pada specificity untuk mengidentifikasi setiap orang yang mempunyai risiko serius. Tidak ada tes yang akan benar-benar akurat 100% sensitive dan 100% specific. Hal ini penting untuk mengidentifikasi dampak dari kesalahan diagnosis ketika diberikan tes tertentu.

Positive Predictive Value

Perhitungan positive predictive values digambarkan pada tabel 2 x 2 di berikut ini:

	MRI POSITIVE	MRI NEGATIVE	
Lachman test Positive	A True positives n = 39	B False positives n = 10	PPV A/A + B 39/39 + 10 = 0.80
Lachman test Negative	C False negatives n = 8	D True negatives n = 3	NPV D/C + D 3/8 + 3

Tabel 4. Kalkulasi Positive Predictive Value

Positive predictive value (PPV) menggambarkan proporsi (persentase) orang-orang dengan hasil positif pada tes studi yang memiliki masalah seperti yang ditentukan oleh gold standard. PPV memberikan hasil gold standard pada mereka yang positif pada tes yang diuji.

Berikut ini adalah formula PPV:

$$\begin{aligned} & A/(A + B) \\ & 39/(39 + 10) \\ & 39/49 = 0.80 \end{aligned}$$

Dalam contoh di atas, 80% pasien dengan Lachman test positif mengalami kerobekan ACL robek yang dikonfirmasi dengan MRI (positif MRI).

Negative Predictive Value (NPV)

Negative predictive value (NPV) menggambarkan proporsi (persentase) orang-orang dengan hasil negatif dalam tes studi yang tidak memiliki masalah seperti yang ditentukan dengan gold standard. NPV memberikan informasi hasil gold standard untuk orang-orang yang telah diuji negatif dalam tes yang diuji.

Beikut ini adalah formula NPV:

$$\begin{aligned} & D/(C + D) \\ & 3/(8 + 3) \\ & 3/11 = 0.27 \end{aligned}$$

Dalam contoh di atas, 27% pasien dengan hasil Lachman test negative yang juga tidak mengalami kerobekan ACL.

Salah satu keterbatasan utama untuk menggunakan predictive values adalah pengaruh prevalensi masalah kepentingan dalam sampel penelitian. Prevalensi lebih rendah dalam sampel menghasilkan PPV lebih rendah. Kembali ke contoh 60 pasien dalam studi utilitas diagnostik Lachman test untuk menentukan ACL rupture, jika kita mengubah prevalensi ACL rupture dalam sampel dari 39 benar positif untuk positif benar 16, PPV turun menjadi 62%. Agar valid, penelitian harus secara akurat mewakili prevalensi penyakit dalam contoh studi yang diharapkan dalam populasi yang lebih luas. Selain itu, sampel 60 pasien mungkin tidak termasuk persentase yang besar dari orang-orang yang diuji negatif pada Lachman test. Dengan demikian, NPV harus ditafsirkan hati-hati. Baik PPV dan NPV dipengaruhi oleh prevalensi kondisi pada sampel penelitian dan harus diinterpretasikan dengan hati-hati.

Bagian: Ringkasan Clinical Bottom Line pada Studi Diagnostik

Sensitivity dan specificity dapat membantu menjawab pertanyaan, "Apa tes terbaik untuk membantu saya menegaskan atau mengesampingkan masalah tertentu?" tetapi hanya jika nilai memiliki telah dihitung pada sampel yang mencakup seluruh spektrum gangguan atau minimal dalam sampel yang mirip dengan pasien Anda. Serangkaian cut off menggunakan nilai untuk menentukan perbedaan antara hasil tes positif dan negatif. Namun, sensitivity dan specificity yang tidak bermfaat secara statistik untuk dikomunikasikan dengan pasien. Statistik ini merujuk kepada tes, tidak secara langsung kepada pasien.

Pertanyaan 7: Apakah tes secara akurat dan klinis relevan terhadap praktik fisioterapi?

Pertanyaan 8: Apakah hasil akan mempengaruhi penanganan saya terhadap pasien?

Tes diagnostik harus akurat dalam menentukan status diagnostik pasien. Selain itu tes harus relevan dengan tujuan pasien dan tujuan penanganan fisioterapi. Setiap tes harus dipilih tidak hanya aspek akurasi tetapi juga secara klinis bermanfaat terhadap pengambilan keputusan klinis.

Likelihood Ratios

Likelihood ratios mengkombinasikan sensitivity dan specificity suatu tes. Nilai positive likelihood ratio akan menjawab pertanyaan, "Apa kemungkinan

bahwa pasien saya dengan hasil tes positif benar-benar memiliki masalah?" Terkait dengan pertanyaan itu, rasio negatif akan menjawab pertanyaan, "apa yang kemungkinan bahwa pasien saya dengan hasil tes negatif memiliki masalah? Likelihood ratios dapat membantu kepastian diagnostik setelah kita mendapatkan hasil tes. Nilai-nilai ini memberikan informasi kuantitatif yang tentang diagnosis klinis.

Likelihood ratios dapat digunakan untuk mengkomunikasikan hasil tes pasien dengan memberikan perubahan probabilitas bahwa masalah yang ada atau tidak (hasil tes positif atau negatif). Karena itu, riwayat pasien, gejala, dan pengujian diagnostik dapat menggambarkan kemungkinan bahwa mereka memiliki masalah gerakan tertentu. Anda dapat menggunakan pernyataan umum: "Saya percaya sangat mungkin bahwa anda memiliki ketidakstabilan lutut akibat kerobekan ACL."

Positive Likelihood Ratio

Positive likelihood ratio (LR+) menggambarkan nilai numerik pada suatu tes yang positif jika ada persoalan gerak. Perhitungan formula LR+ adalah sebagai berikut:

$$LR+ = 1 - \text{Specificity (true negative rate)}$$

Negative Likelihood Ratio

Negative likelihood ratio (LR-) menggambarkan nilai numerik suatu hasil tes negative jika terdapat persoalan gerak. Formula LR- adalah sebagai berikut:

$$LR- = 1 - \text{Sensitivity (false-negative rate)}$$

Clinical Prediction Rules

Clinical predictions rules (CPR), yang sering disebut dengan clinical decision rules adalah suatu set validasi variabel prediktor berdasarkan karakteristik pasien. CPR dapat digunakan untuk penentuan diagnosis yang bersifat spesifik, penentuan prognosis dengan pertimbangan faktor-faktor prognosis serta waktu yang diperlukan, serta rekomendasi treatment terbaik. Dan juga CPR dijadikan sebagai salah satu pertimbangan untuk memberikan referral kepada profesi lain.

Tahap-tahap pengembangan CPR meliputi:

1. Derivasi/penentuan.
2. Validasi.
3. Impact analysis.

E. Daftar Pustaka

Linda Feters, Julie Tilson, Evidence Based Physical Therapy, F.A Davis Company, Philadelphia, 2012), pp 59-69.

Cook CE, Hegedus EJ. Orthopedic Physical Examination Techniques: An Evidence Based Approach. Upper Saddle River, NJ: Pearson/Prentice Hall; 2008.

Campo M, Shiyko MP, Lichtman SW. Sensitivity and specificity: a review of related statistics and controversies in the context of physical therapist education. J Phys Ther. Educ. 2010;24:69-78.