

Tecnologías Sanitarias Multicriterio

Dra. Karina Aliga Llerena

Instituto Nacional de Enfermedades Neoplásicas

El tema que voy a desarrollar es las evaluaciones de tecnologías sanitarias multicriterio.

Para la inclusión de nuevas terapias, es fundamental entender qué es una evaluación de tecnología sanitaria. Se trata de un proceso multidisciplinario, en el que participan diversas disciplinas, desde financiadores hasta clínicos, metodólogos y pacientes. Se utilizan métodos explícitos, a los que se refiere como evidencia, cómo son las revisiones sistemáticas y ensayos clínicos. El objetivo principal es evaluar o asignar un valor a una tecnología a lo largo de su ciclo de vida, tal como lo ha definido INNATA.

Por otro lado, cuando hablamos de una evaluación de tecnología sanitaria con evaluación multicriterio, nos referimos a la inclusión de múltiples criterios en su análisis. Este valor considera diversos aspectos con el objetivo de mejorar la calidad en la toma de decisiones. El documento técnico nacional lo define como una aproximación para incluir estos múltiples criterios en la evaluación de una tecnología con el fin de tomar una decisión en salud.

Adicionalmente, en la página de RENETSAREDEXA, la Red Latinoamericana de Evaluación de Tecnologías Sanitarias., tienen un portal llamado Brisa donde las agencias miembro publican sus estudios. A través de Brisa, RENETSAREDEXA también compara los perfiles de los países en relación a cómo desarrollan estas evaluaciones, incluyendo países como Argentina, Chile y Brasil.

Es interesante observar cómo estos países aplican el multicriterio, especialmente en aspectos como la calidad de la evidencia, la magnitud del beneficio y el impacto económico.

Finalmente, en cuanto al proceso ETS en nuestro país, consta de cinco fases. En esta ocasión, me centraré en la evaluación de los criterios. El primer paso es formular y ajustar la pregunta clínica, donde participan el equipo metodológico, el representante y el comité farmacoterapéutico.

El comité farmacoterapéutico inicia la solicitud para evaluar esta tecnología, y es el médico especialista tratante quien continúa el proceso. La pregunta PICO, que se deriva de las siglas Población, Intervención, Comparador y Desenlaces, es formulada inicialmente por la institución solicitante. Sin embargo, esta puede ser perfeccionada cuando se incorpora la evaluación del equipo metodológico o cuando otros expertos clínicos se involucran. Puede haber diferencias en definir con precisión la población, especificando más claramente quiénes se beneficiarían del tratamiento, estableciendo un comparador más acorde con la práctica clínica asistencial y definiendo mejor los desenlaces.

Cuando se tienen desenlaces, y me refiero a resultados esperados con la tecnología, como reducción de la mortalidad o disminución de eventos adversos, es necesario graduarlos. La metodología Grey nos indica que debemos categorizarlos como críticos, importantes o de importancia limitada. Los críticos e importantes son esenciales para la toma de decisiones, mientras que los de importancia limitada no se incluyen en el análisis. La calificación de estos desenlaces depende de su validez.

Quiero compartir la definición de un desenlace surrogado o intermedio. En oncología, hay muchos de estos, y generalmente son eventos que no tienen una importancia directa pero que podrían predecir desenlaces finales significativos. Un desenlace final es lo que más le importa al paciente, pero pueden surgir desenlaces intermedios antes de estos, como la supervivencia libre de progresión o la supervivencia libre de metástasis. En nuestro país, se establece que este desenlace intermedio o surrogado debe estar validado, es decir, debe ser biológicamente plausible, predecir la evolución de la enfermedad y traducirse en un efecto en el desenlace final.

Matemáticamente, esto se traduce utilizando los criterios de Edwin, al igual que en Argentina. Un desenlace surrogado es validado cuando se correlaciona fuertemente con el desenlace final. Como ejemplo, en un estudio clínico sobre cáncer de cabeza y cuello localmente avanzado, se intentó determinar si el control locorregional era

un desenlace surrogado de supervivencia global. Según los criterios, se correlaciona cuando el límite inferior del intervalo de confianza es mayor o igual a 0.85.

Es importante señalar que, aunque se está aplicando la validación de desenlaces surrogados, el reglamento menciona la evaluación y adquisición de tecnología sanitaria. En la adquisición, se mencionan estos resultados intermedios, pero no se especifica que deban ser validados, lo que puede generar confusiones. Si estos no resultan ser costo-efectivos, se podrían adquirir a través de mecanismos diferenciados. Es decir, existe una ruta para considerar estos resultados intermedios, siempre que, por supuesto, evidencien eficacia y seguridad.

Otro punto importante son los umbrales clínicos, que se definen previamente a la deliberación final. Un umbral clínico, en términos concretos, es aquella mínima diferencia en la cual se puede manifestar un beneficio adicional. Generalmente, esos desenlaces mínimos corresponden al primer umbral. Aunque se buscan estudios clínicos, en la práctica oncológica, no siempre existen para este tipo de desenlaces mínimos. Cuando no hay, la definición varía dependiendo de la institución encargada de evaluar la tecnología. Por ejemplo, cuando el INS estaba realizando las primeras evaluaciones de tecnologías sanitarias, asumían que esto era apreciado por el grupo de trabajo. Se realizaba una discusión y se llegaba a un consenso. Lo decidía el panel de expertos, tomando como punto de partida lo que se espera del efecto del comparador.

Un ejemplo sencillo es una publicación sobre SORAFENIB en carcinoma hepatocelular. Si se espera que el comparador sea un placebo y que este pueda lograr una supervivencia del 30%, ¿cuánto adicionalmente podría lograr la intervención del SORAFENIB? El panel de expertos podría determinar que un beneficio trivial sería un incremento inferior al 3%, un beneficio pequeño entre el 3% y el 6%, y así sucesivamente.

Es importante mencionar que RENETSARENEXA se ha conformado por tres miembros y que el Ministerio de Salud ha iniciado la evaluación de algunas tecnologías. Por ejemplo, han establecido que determinarán la magnitud del efecto a través de criterios establecidos para la definición de los desenlaces surrogados válidos.

Sin embargo, a diferencia de Argentina, que utiliza la metodología de Edwin (una batería de 5 pruebas consecutivas para el diagnóstico de la afectación del sistema nervioso autónomo o vegetativo) tal como está, en nuestro país se ha hecho una modificación.

Finalmente, se establecen intervalos de confianza para cada desenlace, y dependiendo de dónde se ubique la tecnología, se definirá si el beneficio que brinda en relación al desenlace de mortalidad es pequeño, moderado o grande.

Ahora, entrando a los criterios, espero no haberlos confundido más. Quiero comentarles que nuestro país, a diferencia de otros que mencioné, considera más criterios para la toma de decisiones. Esto es positivo porque nos permite analizar varios aspectos de la tecnología.

Un clínico podría valorar ciertas características, mientras que un financiador podría considerar otras. En Perú, se utilizan 10 criterios.

El primero es la **carga de enfermedad**. Este no tiene un valor puntual, sino que es referencial. Se considera la incidencia, mortalidad y los años de vida saludable perdidos.

Luego, tenemos la **necesidad clínica**, que en esencia determina si existen otras alternativas en el país. Esto se verifica en el Petitorio Nacional Único de Medicamentos o la lista complementaria. Sin embargo, es esencial que la tecnología comparativa esté en uso habitual en la práctica clínica, por lo que debe correlacionarse con guías locales o internacionales.

Otro criterio son los **efectos deseables, es decir, la eficacia**. Se valora la magnitud de los beneficios adicionales de la tecnología sanitaria en desenlaces como supervivencia global, calidad de vida y supervivencia libre de progresión. Estos umbrales pueden ser moderados, grandes, pequeños, triviales, variables o indefinidos. Evidentemente, un desenlace con una eficacia adicional moderada o grande es preferible.

A modo de ejemplo, se presenta una tabla que muestra los datos relacionados con NIVOLUMAB para pacientes con melanoma maligno y resecable. En esta tabla, se observa la eficacia en el desenlace de supervivencia global, la intervención, el comparador, el efecto relativo y otros datos relevantes. NIVOLUMAB logra una reducción de la

mortalidad del 27%, mientras que con la quimioterapia, la mortalidad alcanza casi el 60%. Esta diferencia del 27% representa el beneficio adicional que obtenemos con la inclusión de este medicamento.

Otro desenlace de eficacia que se mide es la **calidad de vida**. Aquí, un mayor puntaje indica una mejor calidad de vida. Con la intervención de NIVOLUMAB, se observa un aumento de 1.8 puntos, mientras que con la quimioterapia, es de 0.9 puntos. Esa diferencia de 0.9 puntos es el beneficio adicional que proporciona NIVOLUMAB en términos de mejora de la calidad de vida.

Finalmente, los efectos indeseables están relacionados con la **seguridad** y se busca valorar la magnitud de los daños anticipados con la tecnología sanitaria. Los desenlaces de seguridad que se consideran incluyen eventos adversos graves y otros eventos adversos.

Los **efectos indeseables**, por ejemplo, aquí, son serios. La intervención producía un 20% de eventos adversos, y el comparador que utilizamos, un 26%. Esta diferencia es casi del 7%. Entonces, si lo vemos en números, podríamos evitar, por ejemplo, 71 eventos adversos serios por cada 1,000 pacientes si incluyéramos NIVOLUMAB dentro de nuestro arsenal terapéutico.

Otro criterio que utilizamos es la **certeza y la evidencia**. Todos los países de la región ya utilizan la metodología GRADE (método para valorar la certeza de la evidencia y para la formulación de recomendaciones que ayuden en la toma de decisiones en salud), que se utiliza ampliamente para desarrollar guías de práctica clínica y para realizar estas evaluaciones de tecnologías sanitarias.

Es la misma metodología, sin ninguna variación, que clasifica la evidencia en alta, moderada, baja o muy baja. Tener una certeza alta o moderada significa que, obviamente, puedo confiar en los resultados que veo al introducir la molécula. Por ejemplo, tanto el INS como otra institución (que creo que es de Argentina) han catalogado un ensayo clínico con la misma certeza de evidencia.

El siguiente criterio es el **balance de efectos**, es decir, ponderar entre los efectos deseables e indeseables para determinar hacia dónde se inclina la tecnología. Puede favorecer la intervención o, por el contrario, los resultados pueden favorecer al comparador. Este criterio es el nivel de innovación, que en nuestro país se define en base a dos aspectos: la certeza de la evidencia debe ser al menos moderada y el beneficio debe ser sustancial o mayor.

La equidad, se refiere al impacto en la equidad en salud al introducir esta tecnología. Puede variar según el criterio GRADE. Hay diferentes rangos que el panel, basado en una revisión o análisis de evidencia, podría definir. Si hay estudios de equidad, se consideran; si no los hay, es a criterio del panel participante.

En cuanto a los **recursos necesarios**, se hace una estimación de la diferencia de costos. Por ejemplo, para BEVACIZUMAB en quimioterapia para cáncer colorrectal, se estimó el costo de la intervención con quimioterapia estándar (FOLFOX) y se encontró que, dependiendo del escenario, la diferencia es de aproximadamente 70,000 soles. Es decir, se necesitarían 70,000 soles adicionales para incluir este anticuerpo monoclonal en la terapia. Según lo que está aplicando DIGEMID actualmente, esta diferencia debería ser menor a un umbral que representa 41,400 o moderada entre uno y dos. En este caso, BEVACIZUMAB calzaría para costos moderados. Sin embargo, cuando el INS lo evaluó inicialmente, se consideró como costos extensos. Esto muestra que, dependiendo de quién evalúe, puede haber diferencias en los resultados, y es algo que deberíamos evitar.

El costo-efectividad, lamentablemente, es un criterio que no se está desarrollando porque necesitamos estudios de evaluaciones económicas. A día de hoy, la mayoría de los estudios no se incluyen, por lo que es un criterio que, lamentablemente, se pierde dentro del análisis.

Para terminar, les muestro cómo queda el marco final de valoración de los criterios que les he establecido por tecnología. Aquí no hay una ponderación; no significa que a mayor puntaje se incline hacia algo a favor. Todo depende de la discusión del panel. Debe haber consenso y mayoría para determinar qué valora cada uno y poder decidir qué tecnología se incluye a nivel nacional y cuál no.

Hasta la fecha, RENETSARENEXA ha desarrollado 20 tecnologías sanitarias, de las cuales 11 ya están publicadas. 11 tienen dictamen a favor y 9 en contra. Puede ser que estén en contra porque no hay evidencia clínica, no aumentaron la equidad, la certeza fue muy baja, los beneficios fueron mínimos, no fueron coste-efectivos, no innovadores, o una combinación de estos.

Hay medicamentos que, tras la evaluación, ya no continúan siendo administrados a los pacientes. Estos son los que tienen este dictamen.

En cuanto a los retos, hay muchos. La demanda de solicitudes comenzó con 28 a cargo del INEM, y hoy ya son 103 entre todos los comités farmacoterapéuticos a nivel nacional. Aunque esto no me asusta, ya que en Brasil, por ejemplo, tienen 900 solicitudes. Esto significa que debemos priorizar. Aunque no hay un proceso de priorización establecido, sabemos quién debería hacerlo. Según el documento técnico, la priorización le compete a la Red Nacional.

En promedio, se han demorado 4 meses entre la solicitud y la publicación, lo cual no es mucho comparado con otros que se demoran entre 6 meses a un año.

Otro reto para RENETSARENEXA es que actualmente solo se enfoca en tratamientos. Las DTS abarcan más que solo medicamentos; incluyen promoción, prevención, diagnóstico y el manejo integral del cáncer. Además de medicamentos, las tecnologías pueden ser dispositivos, equipos o procedimientos. En un futuro cercano, deberíamos comenzar a evaluar estos también.

Como ya mencioné, no hay evaluaciones económicas ni documentos técnicos para saber cómo realizarlas. Aunque hay documentos bien hechos en otras regiones, sería cuestión de ver qué modelos externos aplicarían para nuestro país.

Las enfermedades transmisibles y no transmisibles también requieren DTS, y actualmente solo nos estamos enfocando en enfermedades raras y oncológicas.

Finalmente, lo que quiero destacar es que falta consenso en RENETSARENEXA. Aunque está conformada, siento que falta integración para que, de manera unánime, nos indiquen cómo se evaluarán los umbrales clínicos y la validación de los enlaces surrogados. Es esencial que todas las tecnologías sigan las mismas reglas y no dependan de quién las evalúe.

Y, finalmente, la participación formal de los pacientes aún no está ocurriendo. Al inicio del proceso, el INS realizaba consultas a los pacientes, pero con el tiempo, esto se ha ido retirando. Por lo tanto, es esencial trabajar con los pacientes y sus representantes para que también participen en este proceso de evaluación de tecnologías sanitarias.

Gracias.