|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Sección y Cuestión** | **No** | **Ítem** | **Reportado en página #** |
|  |  |  |  |  |
|  | **TÍTULO O RESUMEN** |  |  |  |
|  |  | **1** | Identificación como estudio de precisión diagnóstica, utilizando al menos una medida de precisión (sensibilidad, especificidad, valores predictivos, o AUC) |  |
|  | **RESUMEN** |  |  |  |
|  |  | **2** | Resumen estructurado del diseño, métodos, resultados y conclusiones del estudio (para orientación específica, consúltese la sección de Resúmenes de STARD) |  |
|  | **INTRODUCCION** |  |  |  |
|  |  | **3** | Antecedentes científicos y clínicos, incluyendo el uso previsto y el rol clínico de la prueba de índice  |  |
|  |  | **4** | Objetivos e hipótesis del estudio  |  |
|  | **MÉTODOS** |  |  |  |
|  | *Diseño del estudio* | **5** | Detállese si la recopilación de datos se planificó antes de la prueba de índices, y si se realizaron estándares de referencia previos (estudio prospectivo) o posteriores (estudio retrospectivo) |  |
|  | *Participantes* | **6** | Criterios de elegibilidad |  |
|  |  | **7** | Sobre qué base se identificaron los participantes potencialmente elegibles (síntomas, resultados de pruebas previas, inclusión en el registro) |  |
|  |  | **8** | Dónde y cuándo se identificaron los participantes potencialmente elegibles (centro, localidad y fechas) |  |
|  |  | **9** | Detállese si los participantes formaron series consecutivas, aleatorias o a conveniencia |  |
|  | *Métodos de la prueba* | **10a** | Prueba de índices, suficientemente detallada para permitir la replicación |  |
|  |  | **10b** | Estándar de referencia, suficientemente detallado para permitir la replicación |  |
|  |  | **11** | Fundamentos para elegir el estándar de referencia (si existen alternativas) |  |
|  |  | **12a** | Definición y fundamentos de los puntos de corte de positividad de la prueba o categorías del resultado de la prueba de índices, distinguiendo entre pre-especificado y exploratorio |  |
|  |  | **12b** | Definición y fundamento de los puntos de corte de positividad de la prueba o categorías del resultado del estándar de referencia, distinguiendo entre pre-especificado y exploratorio |  |
|  |  | **13a** | Detállese si los resultados de la información clínica y los estándares de referencia fueron suministrados a los elaboradores/lectores de la prueba de índices |  |
|  |  | **13b** | Detállese si los resultados de la información clínica y de la prueba de índices fueron suministrados a los asesores del estándar de referencia |  |
|  | *Análisis* | **14** | Métodos de cálculo o comparación de las medidas de precisión diagnóstica |  |
|  |  | **15** | Modo de manejo de los resultados indeterminados de la prueba de índices o estándares de referencia |  |
|  |  | **16** | Modo de manejo de los datos ausentes de la prueba de índices o estándares de referencia |  |
|  |  | **17** | Análisis de variabilidad de la precisión diagnóstica, distinguiendo entre pre-especificado y exploratorio |  |
|  |  | **18** | Tamaño de la muestra previsto, y modo de determinarse  |  |
|  | **RESULTADOS** |  |  |  |
|  | *Participantes* | **19** | Flujo de participantes, utilizando un diagrama |  |
|  |  | **20** | Características demográficas y clínicas basales de los participantes |  |
|  |  | **21a** | Distribución de la gravedad de la enfermedad en aquellos sujetos con situación diana |  |
|  |  | **21b** | Distribución de los diagnósticos alternativos en aquellos sujetos con situación diana |  |
|  |  | **22** | Intervalo temporal e intervenciones clínicas entre la prueba de índices y el estándar de referencia |  |
|  | *Resultados de la prueba* | **23** | Tabulación cruzada de los resultados de la prueba de índices (o su distribución)y los resultados del estándar de referencia |  |
|  |  | **24** | Cálculos de precisión diagnóstica y exactitud (95% de intervalo de confianza) |  |
|  |  | **25** | Episodios adversos desde la realización de la prueba de índices y el estándar de referencia |  |
|  | **DISCUSIÓN** |  |  |  |
|  |  | **26** | Limitaciones del estudio, incluyendo fuentes de sesgo potencial, incertidumbre estadística, y generalizabilidad |  |
|  |  | **27** | Implicaciones para la práctica, incluyendo uso previsto y rol clínico de la prueba de índices |  |
|  | **OTRA INFORMACION** |  |  |  |
|  |  | **28** | Número y nombre de registro  |  |
|  |  | **29** | Lugar de acceso al protocolo del estudio  |  |
|  |  | **30** | Fuentes de financiación y otros respaldos; rol de los financiadores |  |
|  |  |  |  |  |

STARD 2015

### OBJETIVO

STARD es el acrónimo de "Standards for Reporting Diagnostic accuracy studies". Este listado de ítems fue desarrollado para contribuir a la integridad y transparencia en el reporte de estudios de precisión diagnóstica. Los autores pueden utilizar dicho listado para redactar informes de estudios informativos. Los editores y pares revisores pueden utilizarlo para evaluar si se ha incluido la información en los documentos suministrados para publicación.

### ExplICACIÓn

Un estudio de precisión diagnóstica evalúa la capacidad que tienen una o más pruebas médicas para clasificar correctamente los participantes del estudio, en su calidad de poseedores de una condición diana. Esta última puede ser una enfermedad, un estadio de la enfermedad, una respuesta o beneficio de una terapia, o un episodio o situación futuros. Una prueba médica puede ser un procedimiento de imagen, una prueba de laboratorio, los elementos de una historia médica y una valoración física, así como una combinación de los mismos, o cualquier otro método recopilatorio de información relativa a la situación sanitaria actual de un paciente.

La prueba cuya precisión se evalúa es denominada prueba de índices. Un estudio puede evaluar la precisión de una o más pruebas de índices. La evaluación de la capacidad de una prueba médica para clasificar correctamente a los pacientes se realiza normalmente comparando la distribución de los resultados de la prueba de índices y los del estándar de referencia. Dicho estándar de referencia constituye el método mejor disponible para establecer la presencia o ausencia de la situación diana. Un estudio de precisión puede basarse en uno o más estándares de referencia.

Si los resultados de la prueba se clasifican como positivos o negativos, la tabulación cruzada de los resultados de la prueba de índices frente a los resultados del estándar de referencia puede utilizarse para calcular la sensibilidad de la prueba de índices (proporción de participantes con situación diana que tienen una prueba de índices positiva), así como su especificidad (proporción de participantes sin situación diana que tienen una prueba de índices negativa). A partir de esta tabulación cruzada (a veces denominada tabla de contingencia, o tabla "2x2"), pueden calcularse otras estadísticas precisas, tales como los valores predictivos positivos o negativos de la prueba. Pueden calcularse entonces los intervalos de confianza que rodean a los cálculos de precisión, para cuantificar la **precisión** estadística de las medidas.

Si los resultados de la prueba de índices pueden adoptar más de dos valores, la categorización de los resultados de la prueba como positivos o negativos requiere un **punto de corte de positividad de la prueba**. Cuando se definen múltiples puntos de corte, los autores pueden reportar una curva de rendimiento diagnóstico (ROC), que representa gráficamente la combinación de la sensibilidad y especificidad para cada punto de corte de positividad de la prueba. El área bajo la curva ROC aporta información, mediante un único valor numérico, acerca de la precisión diagnóstica global de la prueba de índices.

El uso previsto de una prueba médica puede ser el diagnóstico, cribado, estadificación, monitorización, vigilancia, predicción o pronóstico. El rol clínico de una prueba explica su posición con relación a las pruebas existentes en la vía clínica. Una prueba de sustitución, por ejemplo, sustituye a una prueba existente. Una prueba de triaje se utiliza con anterioridad a una prueba existente, y una prueba añadida se utiliza con posterioridad a una prueba existente.

Además de la precisión diagnóstica, existen otros resultados y estadísticas adicionales que pueden ser relevantes para la evaluación de las pruebas médicas. Las pruebas médicas pueden utilizarse asimismo para clasificar a los pacientes para fines diferentes al diagnóstico, como es el caso de la estadificación y el pronóstico. El listado STARD no ha sido explícitamente desarrollado para estos resultados, estadísticas y tipos de estudio adicionales, aunque la mayoría de los ítems de STARD podrían ser aplicados.

### DESARROLLO

Este listado STARD fue elaborado en 2015. Los 30 ítems fueron identificados por un grupo experto internacional de metodólogos, investigadores y editores. El principio orientador para el desarrollo de STARD fue la selección de los ítems que, al reportarse, pudieran ayudar a los lectores a juzgar los sesgos potenciales del estudio, valorar la aplicabilidad de los hallazgos del estudio, y la validez de las conclusiones y recomendaciones. El listado supone una actualización de la primera versión, que fue publicada en 2003.

Para más información visiten la página [http://www.equator-network.org/reporting-guidelines/stard](http://www.equator-network.org/reporting-guidelines/stard/).