

# Diseño de un cuestionario eficaz para el cuidado de heridas

**Palabras clave** análisis, medidas, diseño del cuestionario, tasa de respuesta, validación

**Como referencia** Stephenson J. Diseño de un cuestionario eficaz en el cuidado de heridas . Revista WCET<sup>®</sup> 2022;42(2):24-29

**DOI** <https://doi.org/10.33235/wcet.42.2.24-29>

Presentado el 29 de abril de 2022, aceptado el 6 de mayo de 2021

## INTRODUCCIÓN

La recogida de datos cuantitativos a través de un cuestionario es una práctica habitual en el cuidado de heridas. Los cuestionarios son una forma relativamente barata y rápida de recopilar datos, y no requieren necesariamente que el investigador esté presente mientras se recogen los datos. Muy a menudo son la única forma viable de recoger los datos necesarios. Los usos habituales de los cuestionarios en el cuidado de heridas, que pueden incluir cuestionarios administrados al personal clínico, a los pacientes o a ambos, incluyen:

- Evaluar la eficacia de un programa de formación clínica para aumentar los conocimientos del personal sobre una determinada enfermedad.
- Evaluar el grado de utilización de un apósito concreto en un determinado entorno clínico.
- Para evaluar un nuevo equipo.
- Para controlar la cicatrización de las heridas bajo un nuevo régimen de tratamiento.
- Evaluar un resultado relacionado con el paciente, como el dolor, la calidad de vida o la satisfacción con el tratamiento recibido.

Aunque existen muchos cuestionarios totalmente validados "fuera de la norma", los investigadores en el ámbito del cuidado de heridas pueden descubrir que las medidas específicas que recogen estos cuestionarios no se ajustan a los objetivos de su estudio propuesto, por lo que puede ser necesario diseñar un instrumento a medida. La investigación basada en un cuestionario implica una cuidadosa reflexión sobre la selección de la muestra del estudio, la maximización de la tasa de respuesta, la identificación de las medidas que se van a evaluar, la formulación y puntuación de los ítems que lo componen, el encuadre de los ítems para el análisis, la consideración de las medidas de resultado y la puntuación de los ítems, y el pilotaje del cuestionario.

## ¿A QUIÉN HAY QUE ENTREGAR EL CUESTIONARIO?

El concepto de generalización -la capacidad de inferir más

allá de los datos de la muestra (los que han rellenado el cuestionario) a una población matriz normalmente mucho más amplia- es clave para la mayoría de los estudios de investigación cuantitativa. Esto requiere una muestra representativa de encuestados. Es casi imposible crear una muestra que refleje exactamente a la población que se supone que representa en todos los aspectos. Se necesitan conocimientos clínicos para establecer rasgos importantes -como el nivel de trabajo, la comorbilidad del paciente o el tipo de herida- que variarán de un estudio a otro. Para determinar si una muestra refleja realmente a la población matriz en las características que se consideran más importantes para el estudio, puede ser necesario conocer al menos la distribución aproximada de las categorías de unidades en la población de interés: por ejemplo, se puede conocer la composición de un equipo típico de enfermería de viabilidad tisular en una organización típica, y los investigadores pueden tratar de reflejar esa composición en el personal invitado a rellenar nuestro cuestionario. Si no se garantiza que la muestra no difiere en algún aspecto importante de la población a la que pretende representar, se puede producir un sesgo de selección que puede debilitar o invalidar los resultados.

Algunas características específicas se aplican a los datos recogidos en muchos estudios sobre el cuidado de heridas. En primer lugar, a menudo hay que recoger datos simultáneamente sobre el personal clínico y los pacientes. Un ejemplo podría ser un estudio de la carga de trabajo de un equipo de enfermería comunitaria en el que se encuestará tanto a los enfermeros como a sus pacientes; normalmente, se aplicarán diferentes conjuntos de cuestionarios a los enfermeros y a los pacientes. Esto suele dar lugar a datos agrupados, en los que un miembro del personal tratará a varios pacientes. En segundo lugar, la unidad de análisis en los estudios sobre el cuidado de heridas no es siempre una persona individual, como suele ocurrir en otras ramas de las ciencias clínicas. Puede tratarse de una herida, como una lesión por presión, y un paciente puede aportar varias heridas al mismo estudio. Una vez más, esto nos lleva a la cuestión de la agrupación de los datos; en este caso, con las lesiones por presión agrupadas dentro de los pacientes individuales.

## MAXIMIZAR EL ÍNDICE DE RESPUESTA

La recogida de datos a través de un cuestionario es especialmente susceptible al sesgo de respuesta, un sesgo introducido por las diferencias de características entre los que deciden completar el cuestionario y los que no. Aunque existen métodos informáticos para imputar los valores de los

### John Stephenson

PHD FRSS(GradStat) CMath(MIMA)

Profesor titular de estadística biomédica

Universidad de Huddersfield, Reino Unido

Correo electrónico [J.Stephenson@hud.ac.uk](mailto:J.Stephenson@hud.ac.uk)

datos que faltan, estos métodos pueden no ser viables en todas las situaciones y, por lo general, es preferible maximizar tanto la proporción de encuestados potenciales que realmente responden, como la proporción de los que responden que dan un conjunto completo de respuestas. Las bajas tasas de respuesta también reducen la potencia del análisis, es decir, la capacidad de detectar cualquier efecto que pueda existir.

Hay algunos métodos obvios para aumentar las tasas de respuesta y de finalización:

- Utilización de formatos electrónicos en lugar de, o además de, cuestionarios en papel (se pueden enviar recordatorios amables por correo electrónico a los que no respondan a intervalos adecuados).
- Evitar los cuestionarios con excesivos items. Todos los items incluidos deben tener un propósito específico: cada item superfluo aumenta la posibilidad de que un encuestado no complete el cuestionario correctamente. Por ejemplo, no se debe pedir a los encuestados que proporcionen directamente información sobre cantidades como el IMC, que pueden ser calculadas por los investigadores a partir de otra información proporcionada por los encuestados.
- Evitar los artículos de redacción ambigua. Los items deben ser rápidos de responder para los encuestados, ofreciendo una selección de opciones o escalas analógicas visuales en lugar de pedir texto libre. El suministro de items condicionales puede introducir confusión y debe limitarse.
- Garantizar el anonimato de los participantes, si es apropiado para la información recogida.

Algunos estudios requerirán que se recojan datos basados en cuestionarios en múltiples ocasiones, por ejemplo, para controlar la calidad de vida o el dolor en pacientes con heridas crónicas. Un problema común en este caso es que la proporción de cuestionarios completados suele disminuir en cada punto de recogida de datos. Esto puede introducir un sesgo adicional en forma de sesgo de deserción, cuando los que se pierden en el seguimiento son de alguna manera sistemáticamente diferentes de los que devuelven los cuestionarios. Aunque no se puede hacer nada con respecto a los pacientes que se marchan o mueren durante el periodo de seguimiento, se puede minimizar la pérdida por desgaste si no se sobrecarga a los encuestados en cuanto a la frecuencia de los envíos del cuestionario, ni a la longitud o complejidad de los cuestionarios que deben rellenar.

## VALIDACIÓN / MEDIDAS A EVALUAR

La elaboración de items apropiados para encapsular eficazmente las medidas de resultado de interés suele ser la parte más difícil del diseño de un cuestionario eficaz. Por lo general, es preferible utilizar un cuestionario validado para su aplicación en participantes similares. Sin embargo, la validación completa es un proceso extenso: Price y Harding<sup>1</sup> informaron del desarrollo y la validación de un cuestionario para medir el impacto de las heridas crónicas (úlceras en las piernas y úlceras en el pie diabético) en la calidad de vida relacionada con la salud (HRQoL) del paciente e identificar las áreas de preocupación del paciente. El proceso consistió en tres etapas: un grupo de discusión y una serie de entrevistas semiestructuradas para generar los items del cuestionario; un proceso piloto del cuestionario con análisis de los datos

mediante un análisis factorial; y la evaluación de la fiabilidad, validez y reproducibilidad de la escala resultante en un periodo de seguimiento de 3 meses.

Aunque la validación completa de un cuestionario diseñado por uno mismo es una empresa importante que puede no estar dentro de los recursos de un clínico que necesita diseñar, aplicar y analizar los datos en un periodo de tiempo limitado, algunos pasos comunes de validación pueden ser plausibles. A menudo, esto implicará la contribución a la redacción de los items por parte de un panel de clínicos expertos, con la claridad de la redacción posiblemente evaluada a través de grupos de discusión u otros medios. El objetivo es obtener una serie de items que contribuyan cada uno a una faceta diferente del resultado de interés y que, evaluados conjuntamente, proporcionen una medida significativa del resultado global. Puede ser necesario el asesoramiento de un experto para confirmar que un item contribuye realmente a la medición del constructo previsto, y no de otro constructo. Barakat-Johnson et al<sup>2</sup> desarrollaron y evaluaron las propiedades psicométricas de un instrumento utilizado para evaluar los conocimientos de los clínicos sobre la dermatitis asociada a la incontinencia, con el desarrollo de los items utilizando la aportación de un panel de expertos de clínicos como primera etapa de un proceso de tres fases; a esto le siguió una evaluación de la validez del contenido del instrumento a través de una encuesta a los clínicos y a las partes interesadas, y un diseño piloto de encuesta transversal en varios sitios para determinar la fiabilidad del compuesto.

La validez de contenido y de constructo también debe abordarse durante el proceso de desarrollo. Deben evitarse los items demasiado autosimilares. En lugar de que cada uno de los items capte una faceta única del constructo de interés, dichos items están captando la misma faceta y, por lo tanto, esta faceta se está contando doblemente, y es muy probable que los encuestados respondan de la misma manera a ambos items. Sin embargo, a la inversa, los items que son muy diferentes entre sí pueden no estar midiendo el mismo constructo en absoluto. Otro problema común es el "solapamiento" de las facetas de un constructo captadas por diferentes items. Barakat-Johnson et al.<sup>2</sup> utilizaron la evaluación de la validez de contenido y de constructo mediante medidas de resumen y métodos estadísticos reconocidos en las fases posteriores de desarrollo de su herramienta.

## FORMULACIÓN Y PUNTUACIÓN DE LOS ITEMS

La obtención de datos cuantitativos a través de un cuestionario requiere respuestas "cerradas" (números o categorías); las respuestas "abiertas" no suelen ser adecuadas para los informes cuantitativos. Los items del cuestionario de forma cerrada pueden formularse de varias maneras. Algunas de las formulaciones de artículos más comunes son:

- Items que solicitan directamente una cantidad numérica, como "¿Cuál es su edad en años?"
- Items que arrojan una cantidad numérica de forma indirecta, solicitando a los encuestados que den una respuesta en una escala analógica visual que posteriormente es procesada por el investigador. Un ejemplo típico podría ser presentar una línea de una longitud determinada (digamos 10 cm) con ambos

extremos claramente etiquetados como valores extremos; por ejemplo: *"Ningún dolor"* y *"El peor dolor imaginable"*; y acompañados de una instrucción como *"Por favor, ponga una marca en esta línea correspondiente al nivel de dolor que le causa su herida hoy"*.

- Items que permiten a los encuestados elegir una opción de una lista de posibles opciones ofrecidas.
- Items que permiten a los encuestados elegir tantas opciones como sean aplicables de una lista de posibles opciones ofrecidas.

Los dos primeros tipos de respuestas son numéricos y los dos segundos son categóricos. Ambos tipos de respuestas pueden ser potencialmente útiles para el análisis posterior, y el cuestionario debe estar formateado de tal manera que sea posible que los encuestados informen de una respuesta numérica, o que elijan de una lista de opciones, según sea el caso, a un ítem en particular.

Los ítems que provocan respuestas numéricas directas o indirectas son potencialmente los más sencillos de incluir en los procedimientos de análisis posteriores. Sin embargo, el preprocesamiento posterior de los datos puede ser más fácil si se formula una pregunta de tal manera que los encuestados no sientan la necesidad de añadir palabras innecesarias: una pregunta como *"¿Cuánto tiempo ha trabajado en esta organización?"* puede suscitar una serie de respuestas como *"Menos de 1 año"*; *"18 meses"*; *"Unos 5 años"*, etc., que la mayoría de los programas informáticos interpretan como respuestas textuales, en lugar de numéricas, y que requieren una amplia edición antes de poder utilizarlas para el análisis. Un simple cambio de redacción como *"Indique el número de años (redondeado al año más cercano) que ha trabajado para esta organización"* podría ahorrar mucho tiempo de preprocesamiento. Además, una simple instrucción de dejar en blanco los ítems no aplicables, o los ítems para los que el encuestado no puede dar una respuesta correcta, puede ahorrar más tiempo al eliminar varios casos de *"No aplicable"*; *"No sabe"*; *"No está seguro"*, etc.

Es una práctica habitual introducir una categorización artificial en los ítems que arrojan datos numéricos. Por ejemplo, un ítem en el que se pide a los encuestados que informen sobre su edad podría ofrecer una opción de rango de edad: *"18-30"*, *"31-40"*, *"41-50"*, etc. Por lo general, no se recomiendan estos enfoques: en primer lugar, se pierde información sobre la distinción entre los encuestados de diferentes edades dentro del mismo rango de edad (puede haber diferencias considerables en las respuestas de una persona de 18 años y las de una de 30 años); y en segundo lugar, las múltiples categorías en una variable de agrupación significan que se necesitan múltiples comparaciones en el análisis (resultados en los de 18-30 años frente a los de 31-40 años, resultados en los de 18-30 años frente a los de 41-50 años, etc.), lo que puede dar lugar a problemas técnicos y de interpretación.

Sin embargo, en el caso de los ítems que captan un constructo realmente medido a nivel categórico, no hay alternativa a ofrecer una lista de opciones para que los encuestados las seleccionen. La lista de opciones ofrecidas debe ser exhaustiva. Un encuestado al que se le pida que indique su función en una organización, por ejemplo, sólo para descubrir que su función no está representada en las opciones ofrecidas, puede perder la

confianza en que su participación en el estudio dará lugar a un registro exacto de su opinión o situación y puede estar menos dispuesto a completar el resto del cuestionario con exactitud.

Un problema similar se plantea cuando las opciones se solapan. Si las opciones de la pregunta *"¿Cuántos pacientes tiene por semana?"* son, por ejemplo, *"10 o menos"*; *"10-20"*; *"20-30"*, etc., entonces alguien con una carga de trabajo de 10 o 20 pacientes exactamente no sabrá qué opción debe seleccionar. Otro ejemplo podría ser un encuestado al que se le pide que seleccione su función laboral de una lista de opciones cuando en realidad tiene dos o más funciones. Esta situación puede evitarse simplemente con una mejor redacción del artículo, por ejemplo: por favor, seleccione la función de la siguiente lista que más se aproxime a su función principal".

Al formular ítems de este tipo, puede ser tentador permitir a los encuestados una respuesta de texto libre. Esto puede evitar la omisión accidental de la opción preferida por el encuestado, o la confusión derivada de múltiples opciones que son similares, pero no idénticas, a la respuesta que el encuestado preferiría dar. Sin embargo, esta posibilidad puede requerir un amplio preprocesamiento posterior de los datos de texto libre en grupos definidos, lo que no siempre es fácil si los encuestados no son lo suficientemente explícitos en sus respuestas de texto libre. Esta situación puede evitarse a menudo ofreciendo una opción "Otros" en la lista de opciones.

Las opciones ofrecidas a un ítem categórico pueden ser nominales (sin ordenación subyacente; en cuyo caso la ordenación de las opciones no es importante) u ordinales (en cuyo caso las opciones deben presentarse en un orden lógico). El ítem ordinal "clásico" del cuestionario es el ítem Likert, la formulación más simple y, por cierto, la más popular de los ítems del cuestionario, que se encuentra en muchos, si no en la mayoría, de los cuestionarios. Un ítem de Likert es una pregunta que suele pedir a los encuestados que elijan una opción de una lista ordenada de cinco opciones que representan el grado de acuerdo con una determinada afirmación, como, por ejemplo, *"El producto X es un tratamiento eficaz para la sobregulación"*. Las opciones típicas de un ítem de este tipo podrían ser *"Muy en desacuerdo"*, *"En desacuerdo"*, *"Ni de acuerdo ni en desacuerdo"*, *"De acuerdo"* y *"Muy de acuerdo"*. Otros ítems de Likert pueden pedir a los encuestados que evalúen la frecuencia o la magnitud de un acontecimiento, como, por ejemplo, *"¿Se ha hinchado la zona alrededor de la herida?"* En este caso, las opciones típicas pueden ser *"Nada"*, *"Un poco"*, *"Una cantidad moderada"*, *"Bastante"*, *"Mucho"*.

Los ítems Likert no tienen que ofrecer cinco opciones, pero en general ofrecen un número impar de opciones, de las cuales cinco es probablemente el número más común, para permitir una opción intermedia "neutral". Aunque los ítems con mayor número de opciones pueden parecer que ofrecen una mayor granularidad de respuesta, las distinciones entre los puntos de la escala pueden ser cada vez más difíciles de discernir para los encuestados (*"Algunas veces"*, *"Muchas veces"*, *"La mayoría de las veces"*, *"Casi todas las veces"*, etc.). Un equivalente visual del ítem Likert es una pregunta redactada de forma similar: *"En una escala de 0 a 10, ¿en qué medida su herida le ha impedido realizar las tareas domésticas cotidianas?"* Este es un ítem de 11 puntos: un error común es permitir que la escala en preguntas de este tipo vaya de 1 a 10 (en lugar de 0 a 10). La respuesta neutra

en estos casos estaría representada por una respuesta de 5,5, no de 5; aunque muchos de los que responden con el valor 5 a ítems de este tipo estarían sin duda intentando informar de una respuesta en el centro exacto de la escala disponible. Los ítems con un amplio conjunto de respuestas ordinales se comportan en cierto modo como los ítems que dan respuestas numéricas indirectamente a través de una escala analógica visual.

Los ítems que piden a los encuestados que seleccionen "tantas opciones como sean aplicables" son aceptables, pero estos ítems pueden ser significativamente más difíciles de analizar que los ítems correspondientes que piden que se elija una sola opción. Por ejemplo, un ítem como "*¿Cuál de los siguientes apósitos utiliza usted regularmente - por favor, seleccione todos los que correspondan*" seguido de una lista de 26 opciones (Producto A, Producto B, Producto C... Producto Z), equivale en realidad, en términos de análisis, a una serie de 26 preguntas: "*Utiliza regularmente el producto A para apósitos, ¿sí o no?*"... "*Utiliza regularmente el producto B para apósitos, ¿sí o no?*"... "*Utiliza regularmente el producto Z para apósitos, ¿sí o no?*". Esta serie de ítems probablemente dará lugar a una amplia gama de combinaciones de respuestas y dará lugar a docenas de comparaciones por pares, todas las cuales serán difíciles de interpretar.

## ENCUADRAMIENTO DE LOS ÍTEMS DE ANÁLISIS

Un cuestionario típico puede comenzar con algunas preguntas demográficas básicas, para obtener los atributos demográficos y de estilo de vida de los encuestados, como la edad, el sexo, la situación familiar, etc.; y/o ítems relativos a su estado de salud (presencia de diversas condiciones de salud mental o física, duración de la herida preexistente) o situación laboral (duración del servicio, grado del personal, etc.). Algunos de estos ítems pueden incluirse para ayudar a ilustrar la diversidad o las características de la muestra, pero no participarán en el análisis propiamente dicho.

Dentro de lo razonable, los ítems que miden estas "variables de fondo", que suelen ser preguntas fácticas que suscitan respuestas numéricas o categóricas, en lugar de ítems tipo Likert o similares, pueden registrarse de la forma que se desee. Los cuestionarios diseñados para presentar datos de forma descriptiva, pero que no implican ningún tipo de análisis inferencial (es decir, inferir a partir de los datos de la muestra a una población matriz), pueden limitarse a ítems de este tipo. Estos estudios suelen estar diseñados para evaluar la prevalencia o la proporción de una cantidad, como un estudio para determinar la proporción de enfermeros que utilizan un determinado producto para el cuidado de heridas, o la proporción de personal clínico que responde a una indicación visual como el enrojecimiento de la piel. Brown y Sneddon<sup>3</sup> aplicaron un cuestionario, compuesto en su mayor parte por ítems "independientes" con respuestas ordinales, para comprender cómo se financian y prestan los servicios de linfedema en todo el Reino Unido y su nivel de recursos. Los datos del cuestionario permitieron estimar las proporciones (por ejemplo, la proporción de médicos encuestados que trataron heridas abiertas), pero los investigadores no intentaron generalizar más allá de los datos de la muestra.

Sin embargo, el análisis inferencial suele estar dentro del

ámbito de la mayoría de los estudios cuantitativos y, por lo tanto, la mayoría de los cuestionarios que obtienen datos cuantitativos incluirán ítems que son necesarios para el posterior análisis inferencial. Por ejemplo, con respecto a un determinado resultado o resultados, se puede desear comparar al personal experimentado y al novato, o a los pacientes de la UCI que se giran regularmente y a los que no, o a un nuevo equipo y a un equipo estándar. Estos análisis son ejemplos de estudios comparativos, en los que se comparan dos o más grupos entre sí: muchos diseños de estudios de investigación estándar, como los estudios de cohortes, los estudios de casos y controles y los diseños controlados aleatorios, entran en esta categoría. Ousey et al.<sup>4</sup> utilizaron datos basados en cuestionarios para comparar un nuevo diseño de colchón con un colchón estándar en una serie de métricas de la experiencia de los pacientes (comodidad, temperatura y calidad del sueño). Los investigadores utilizaron métodos estadísticos inferenciales estándar para comparar la importancia y la magnitud de los efectos, con grupos definidos por el tipo de colchón.

Los ítems utilizados para definir las variables de agrupación en estos estudios son categóricos. Las variables categóricas que sólo pueden adoptar una de las dos categorías (o "niveles", como a veces se conocen) se conocen como variables binarias, como en el estudio de Ousey et al.<sup>4</sup>. Algunas variables de agrupación pueden comprender más de dos categorías. Por ejemplo, un estudio que compare los resultados en pacientes que pueden clasificarse como de bajo peso, de peso normal, con sobrepeso, con obesidad o con obesidad mórbida, podría utilizar una variable de agrupación "Estado de obesidad" para clasificar a cada encuestado en una de las cinco categorías anteriores.

Estas variables de agrupación multicategórica deben especificarse con precaución; mientras que una variable de agrupación binaria conduce a un único análisis (por ejemplo, resultado en varones frente a resultado en mujeres), el número de análisis necesarios aumenta rápidamente con la introducción de variables de agrupación de varios niveles. Otra razón para limitar las variables de agrupación de niveles múltiples es que, aunque los ítems que registran las variables de agrupación deberían, en general, permitir la selección por parte del encuestado de cualquier ítem posible, los investigadores deberían estar preparados para la eventualidad de que los datos estén poco repartidos entre las múltiples categorías, lo que daría lugar a algunos grupos que son realmente demasiado pequeños para analizarlos de forma significativa. En tales circunstancias, puede ser necesario fusionar ciertas categorías antes del análisis.

## MEDIDAS DE RESULTADO

En la mayoría de los cuestionarios, la mayoría de los ítems se refieren a la obtención de medidas de resultado. Muchos resultados son categóricos, a menudo binarios, por ejemplo, la probabilidad de que una herida alcance un 50% de curación a los 30 días del tratamiento; o multicategóricos, por ejemplo, el tipo de tejido predominante en el lecho de la herida. Por lo general, estos resultados pueden recogerse fácilmente en un cuestionario con un único ítem binario u ordinal. Dhoonmoon<sup>5</sup> encuestó la experiencia de 56 profesionales sanitarios (HCP) sobre el uso de una almohadilla de desbridamiento mediante un cuestionario de opinión. La mayoría de los

items, incluidos los relacionados con el rendimiento de las almohadillas (eliminación de los restos de esfacelos, acción de desbridamiento, etc.) se evaluaron mediante items categóricos, con opciones que iban de "excelente" a "deficiente". Estas medidas se prestan naturalmente a la evaluación categórica ordinal. Por ejemplo, uno de los resultados medidos en el estudio de Ousey et al4 (la calidad del sueño) se procesó para el análisis a partir de sus cinco opciones originales ("excelente", "muy bueno", "bueno", "adecuado", "deficiente") en una medida dicotómica que comparaba las respuestas de "excelente" o "muy bueno" con cualquier otra respuesta. También pueden encontrarse resultados numéricos, como el porcentaje de pacientes curados, o el tiempo para que los niveles de dolor alcancen un determinado valor preespecificado, pero son menos comunes en los análisis basados en cuestionarios en el cuidado de heridas.

## PUNTUACIÓN DEL ARTÍCULO

Los cuestionarios suelen utilizarse para evaluar cantidades para las que no existe una medida objetiva sencilla. En el contexto de un estudio sobre el cuidado de heridas, pueden ser, por ejemplo, la evaluación de un clínico sobre un nuevo colchón redistribuidor de la presión, o la opinión de un paciente sobre cuánto le impide su herida realizar las tareas cotidianas. Normalmente, estas cantidades no pueden encapsularse en un solo item; puede ser necesaria una serie de items, todos los cuales se refieren al constructo de interés. Por ejemplo, el conocimiento de la dermatitis de un enfermero en formación que ha realizado recientemente una sesión de taller sobre este tema, o la calidad de vida que experimenta un paciente que vive con una herida crónica. Normalmente, estos items constitutivos pueden ser del tipo Likert o similares. En estos casos, el interés se centra casi siempre en la puntuación procesada de un conjunto de items, y no en ninguno de los items individuales. Por lo tanto, aunque en teoría cada item de un cuestionario podría representar una única medida, el número de medidas distintas recogidas en un cuestionario típico suelen ser mucho menor que el número de items del cuestionario, y varios items contribuyen a la evaluación de cada constructo.

Por lo general, es conveniente limitar el número de resultados: una presentación extensa de los resultados individuales en forma de, por ejemplo, gráficos circulares puede dar poca idea de la importancia relativa de los distintos resultados. También hay ciertos problemas de análisis que pueden hacer que un gran número de resultados primarios no sean deseables. Al igual que los estudios que recogen datos por otros medios, el cuestionario ideal probablemente recoge información sobre un único resultado primario preespecificado y un pequeño número de resultados secundarios.

Se necesita una puntuación para todos los items que contribuyen a la evaluación de una medida concreta. Normalmente, la puntuación de los items Likert de 5 puntos es muy sencilla: de 1 punto para "Muy endesacuerdo" a 5 puntos para "Muy de acuerdo", con las opciones intermedias puntuadas en consecuencia. Los items tipo Likert con otro número de opciones se puntúan de forma similar. Muchos investigadores prefieren utilizar una codificación como: -2 puntos para "Muy en desacuerdo", -1 punto para "En desacuerdo" y así sucesivamente hasta +2 puntos para "Muy de acuerdo", posiblemente con

la idea de que las respuestas redactadas de forma negativa requieren puntuaciones negativas. Esta codificación es exactamente equivalente a la codificación 1-5 mencionada anteriormente: la puntuación de cada opción se reduce en 3 puntos para todas las opciones. Siempre que esta puntuación se aplique de forma coherente, las inferencias serán las mismas con cualquiera de los dos sistemas de puntuación.

Normalmente se asume que las puntuaciones de los items son aditivas, que tiene sentido obtener una puntuación global sumando las puntuaciones obtenidas en los items individuales que contribuyen a la misma medida. Esta suposición suele ser más fácil de justificar si hay coherencia en la formulación de los items. No es obvio cómo debe obtenerse una puntuación global con una serie de items con un número de opciones que varía, por ejemplo, de 2 a 3 y de 5 a 7. Las puntuaciones de los items con mayor número de opciones superarán a las de los items con menos respuestas si, para cada item, las respuestas se codifican simplemente como 1 hasta el valor del número de las opciones.

También es más difícil justificar que la suma de las puntuaciones de varios items conduzca a una medida significativa, aunque el número de opciones de cada item sea el mismo, si las opciones son diferentes. Si un conjunto de items ofrece las opciones "Muy en desacuerdo", "En desacuerdo..." "Muy de acuerdo" y otro conjunto ofrece las opciones "En absoluto", "Un poco..." "Mucho", puede ser difícil argumentar que las puntuaciones de los dos conjuntos de items pueden combinarse de forma significativa.

Para garantizar un total significativo, puede ser necesario invertir la codificación anterior si algunos items van en sentido contrario a otros, por ejemplo, si los items Likert de 5 puntos como "Mi herida me ha obligado a limitar mis actividades con los demás" y "La herida ha afectado a mi sueño" se codifican utilizando la escala de 1 a 5 anterior, con 1 punto otorgado para una respuesta de "Muy en desacuerdo" y 5 puntos otorgados para una respuesta de "Muy de acuerdo", entonces la implicación es que las puntuaciones más altas indican peores resultados. Por lo tanto, si se incluyera un item adicional en la misma escala como, por ejemplo, "Soy capaz de realizar las tareas cotidianas sin dificultad", este item podría codificarse de forma que "Muy de acuerdo" recibiera 5 puntos, "Muy en desacuerdo" 1 punto, y los demás puntos de la escala se puntuaran en consecuencia, por coherencia con el resto de los items de la escala.

## PILOTAJE DEL CUESTIONARIO

La aplicación piloto puede ser una herramienta útil para el perfeccionamiento de los items del cuestionario y puede revelar problemas que pueden repercutir en la tasa de respuesta posterior y en la fiabilidad de las respuestas, como la falta de claridad en la redacción de los items o el tiempo excesivo que se necesita para completar el cuestionario. Si un cuestionario incluye un conjunto de items de tipo Likert o similares que están diseñados para abordar el mismo constructo, la consistencia interna de las respuestas del piloto a estos items puede evaluarse fácil y rápidamente utilizando el software más estadístico. Este proceso puede identificar los items que no se responden de manera similar a otros items que pretenden medir el mismo constructo y, por lo tanto,

pueden requerir modificaciones en su redacción (si la redacción no es clara o ha sido malinterpretada por los encuestados), la eliminación del cuestionario o posiblemente el traslado a la medición de otro constructo. La fase piloto suele ser la única oportunidad para realizar estas modificaciones si son necesarias.

## RESUMEN

Un buen diseño del cuestionario se rige por la pregunta de investigación y el análisis que se deriva de ella. La consideración del punto final es, de hecho, generalmente el punto de partida. Entre las cuestiones que hay que tener en cuenta están la determinación de los resultados que se van a medir; cómo se van a medir; si los resultados son medidas objetivas que se pueden captar adecuadamente utilizando items que provocan respuestas numéricas simples o categorías, o si requieren múltiples items para captar una serie de facetas específicas de la medida.

También hay que determinar el nivel o niveles en los que se va a realizar el análisis: en los estudios sobre el cuidado de heridas, son habituales los análisis a nivel de paciente, de clínico o de herida. También hay que determinar si los resultados se van a relacionar con cualquier otra variable, y si los grupos deseados para la comparación figuran en los items que funcionan como variables de agrupación para clasificar adecuadamente las unidades de análisis (ya sean pacientes, clínicos o heridas).

La recogida de datos a través de un cuestionario debe abordarse de la misma manera que la recogida de datos a través de dispositivos médicos u otros medios: es necesario asegurarse de que el instrumento de recogida de datos es adecuado para su finalidad. Esto significa que se dan tantos pasos como sea posible a lo largo del camino de la validación (suponiendo que no se utilice un instrumento pre-validado) para asegurar que estamos midiendo los resultados que creemos que estamos midiendo, a través de items cuidadosamente redactados, agrupados y puntuados adecuadamente. Hay que tener cuidado de que sólo se utilicen los items necesarios para captar datos demográficos, otra información de fondo y medidas de resultado. Es necesario asegurarse de que los encuestados sean, en la medida de lo posible, una muestra representativa de la población a la que se pretende generalizar. Los índices de respuesta se maximizan haciendo que los items sean lo más claros posible, y pidiendo lo menos posible a los encuestados en cuanto a la duración y el esfuerzo que necesitarán para completar el cuestionario, al igual que podría hacerse con otros medios de recogida de datos.

Aunque es fácil subestimar el esfuerzo necesario para facilitar una recogida de datos eficaz mediante un cuestionario, cuando se lleva a cabo correctamente, la recogida de datos mediante un cuestionario puede ser un medio muy eficaz de recogida de datos y constituir una base sólida para los estudios de investigación.

## CONFLICTO DE INTERESES

Los autores declaran no tener conflictos de intereses.

## FINANCIACIÓN

Los autores no recibieron financiación por este estudio.

## REFERENCIAS

1. Price P, Harding K. Cardiff Wound Impact Schedule: the development of a condition-specific questionnaire to assess health-related quality of life in patients with chronic wounds of the lower limb. *Int Wound J*. 2004 Apr;1(1):10-17.
2. Barakat-Johnson M, Beeckman D, Campbell J, Dunk AM, Lai M, Stephenson J, Coyer F. Development and Psychometric Testing of a Knowledge Instrument on Incontinence-Associated Dermatitis for Clinicians: The Know-IAD. *J Wound Ostomy Continence Nurs*. 2022 Jan-Feb 01;49(1):70-77.
3. Brown L, Sneddon MC. Lymphoedema service provision across the UK: a national survey. *J Lymphoedema*. 2020;15(1):16-21.
4. Ousey K, Stephenson J, Fleming L. Evaluating the Trezzo range of static foam surfaces: results of a comparative study. *Wounds UK* 2016;12(4):66-73.
5. Dhoonmoon L. Experiences of healthcare professionals using Prontosan® debridement pad. *Wounds UK* 2021;17(1):118-123.