



Trabajo Original

Nutrición artificial

Concordancia entre observadores en la aplicación de la lista de verificación para la administración segura de la nutrición enteral *Interobserver agreement in using a checklist for the safe administration of enteral nutrition*

Stella Marys Rigatti Silva^{1,2}, Ana Paula Almeida Corrêa^{1,2}, Michelle Dornelles Santarém^{1,3}, Michelli Cristina Silva de Assis¹ y Mariur Gomes Beghetto²

¹Hospital de Clínicas de Porto Alegre (HCPA), Porto Alegre, Rio Grande do Sul, Brasil. ²Programa de Pós-Graduação em Enfermagem da Universidade Federal do Rio Grande do Sul (PPGENF/UFRGS), Porto Alegre, Rio Grande do Sul, Brasil. ³Escola de Enfermagem da Universidade Federal do Rio Grande do Sul (EENF/UFRGS), Porto Alegre, Rio Grande do Sul, Brasil

Resumen

Objetivo: describir una evaluación de la concordancia entre observadores en el uso de una lista de verificación en la atención a pacientes con sonda nasointestinal (SNE) como premisa de la recogida de datos.

Método: este estudio de fiabilidad se realizó en 2018 en un hospital brasileño, antes de la recogida de datos para un ensayo clínico abierto. Ocho observadores independientes, previamente capacitados y "cegados" con respecto a las evaluaciones de los demás, evaluaron a pacientes con SNE por medio de una lista de comprobación de 25 elementos. Los datos obtenidos por estos asistentes de investigación (AI) se compararon con los obtenidos por una enfermera experimentada. Se midió la concordancia mediante los coeficientes kappa y PABAK.

Resultados: se realizaron 451 observaciones por pares. Considerando la totalidad de los elementos, hubo una concordancia casi perfecta ($k > 0,80$) en todos los pares de observación (enfermera vs. cada AI): AI 1 ($k = 0,91$; IC 95 % = 0,89-0,93); AI 2 ($k = 0,83$; IC 95 % = 0,80-0,85); AI 3 ($k = 0,92$; IC 95 % = 0,90-0,94); AI 4 ($k = 0,83$; IC 95 % = 0,80-0,86); AI 5 ($k = 0,94$; IC 95 % = 0,92-0,96); AI 6 ($k = 0,94$; IC 95 % = 0,92-0,96); AI 7 ($k = 0,96$; IC 95 % = 0,95-0,98); AI 8 ($k = 0,73$; IC 95 % = 0,70-0,77). Se identificó un menor número de concordancia en los elementos individuales y en determinados AI. La capacitación y supervisión de los AI mejoró su desempeño y la concordancia entre observadores.

Conclusión: la evaluación de la concordancia entre observadores resultó fundamental para asegurar la fiabilidad de la recogida de datos y, por consiguiente, evitar sesgos de medición en los estudios clínicos de enfermería.

Abstract

Objective: to describe an evaluation of interobserver agreement in the use of a checklist related to the use of nasoenteral tube (NSS) as a presupposition for quality in obtaining data.

Method: a methodological study conducted in 2018 in a Brazilian hospital, preceding the data collection of an open-label clinical trial. Independent observers, blinded to the evaluation of their peers, evaluated patients with NSS through a 25-item checklist. The data collected by eight previously trained research assistants (RA) were compared to those obtained by an experienced nurse (reference standard). Agreement was measured using the kappa coefficient and PABAK.

Results: a total of 451 observations were made in pairs. Considering the total items on the checklist there was almost perfect agreement ($k > 0.80$) in all observation pairs (nurse vs. each RA): RA 1 ($k = 0.91$; 95 % CI = 0.89-0.93); RA 2 ($k = 0.83$; 95 % CI = 0.80-0.85); RA 3 ($k = 0.92$; 95 % CI = 0.90-0.94); RA 4 ($k = 0.83$; 95 % CI = 0.80-0.86); RA 5 ($k = 0.94$; 95 % CI = 0.92-0.96); RA 6 ($k = 0.94$; 95 % CI = 0.92-0.96); RA 7 ($k = 0.96$; 95 % CI = 0.95-0.98); RA 8 ($k = 0.73$; 95 % CI = 0.70-0.77). However, for isolated items, and in specific RAs, there were fair agreements, unacceptable to effectively collect data from a clinical trial. Retraining and supervision of RAs were able to improve agreement between observers.

Conclusion: an evaluation of interobserver agreement proved to be fundamental to ensure the reliability of data collection and, therefore, to avoid measurement biases.

Palabras clave:

Exactitud de los datos. Variaciones dependientes del observador. Nutrición enteral. Atención de enfermería.

Keywords:

Data accuracy. Observer variation. Enteral nutrition. Nursing care.

Recibido: 15/02/2021 • Aceptado: 17/05/2021

Financiación: el estudio contó con el apoyo del Fondo de Incentivos a la Investigación del Hospital de Clínicas de Porto Alegre (FIPE/HCPA) para la financiación parcial del proyecto.

Conflictos de intereses: los autores informan de la ausencia de cualquier tipo de conflicto de intereses.

Silva SMR, Corrêa APA, Santarém MD, Assis MCS, Beghetto MG. Concordancia entre observadores en la aplicación de la lista de verificación para la administración segura de la nutrición enteral. Nutr Hosp 2021;38(5):903-910

DOI: <http://dx.doi.org/10.20960/nh.03578>

Correspondencia:

Stella Marys Rigatti Silva. Hospital de Clínicas de Porto Alegre (HCPA). Centro de Investigación Clínica, Edificio 21. R. Ramiro Barcelos, 2350 - Santa Cecilia. 90035-007 Porto Alegre, Rio Grande do Sul, Brasil
e-mail: stellamsilva@hcpa.edu.br

INTRODUCCIÓN

La utilización de prácticas seguras puede ayudar a minimizar los eventos adversos asociados a la terapia de nutrición enteral (TNE). Los profesionales de enfermería, responsables de la administración y la supervisión de la TNE (1), garantizan su calidad mediante el cumplimiento de las buenas prácticas y las recomendaciones de las directrices del área.

La búsqueda de estrategias que favorezcan el cumplimiento y la adhesión de los profesionales a las buenas prácticas sigue siendo un desafío para las instituciones hospitalarias. Para mejorar esas actitudes, la inserción de una lista de verificación resultó prometedora para el paciente quirúrgico. Alpendre y sus colegas (2) demostraron, con la validación de la Lista de Seguridad Quirúrgica Pre y Postoperatoria, que esa es una estrategia para promover la seguridad del paciente, evidenciando la importancia de esta herramienta para cambiar el comportamiento de los profesionales de la salud.

En lo que se refiere a la TNE, ningún estudio ha evaluado el uso de una lista de control para la administración segura a través de la SNE. Se sabe que, para la aplicación de los instrumentos, la evaluación de la concordancia entre observadores es un paso que permite establecer su reproducibilidad (3).

Específicamente en la esfera de la nutrición, la evaluación de la concordancia se utiliza para comparar los métodos de evaluación nutricional, los estudios de hábitos alimentarios y la antropometría (4,5). Sin embargo, hay pocos estudios (6,7) que evalúen el consenso entre observadores desde la perspectiva de la seguridad del paciente en el uso de la TNE.

Un estudio realizado por enfermeras (6) evaluó el consenso entre la aceptación comunicada por los pacientes adultos usuarios de la TNE y los registros realizados por las profesionales en las historias clínicas. Se observó que la concordancia entre los informes de los pacientes y los registros de las enfermeras se consideraba de débil a moderada ($\kappa = 0,45$) (6). Otro estudio de concordancia (7) tenía por objeto evaluar el consenso entre el uso de una prueba de auscultación y los rayos X para detectar la posición de la sonda nasointestinal (SNE) (7). Dos enfermeras realizaron la auscultación y dieron su opinión sobre la posición anatómica de la sonda. Hubo una débil concordancia entre las enfermeras (PABAK = 0,054; $p = 0,103$). Además, la concordancia entre los métodos (auscultación y rayos X) era también muy débil para ambas enfermeras (PABAK = 0,188; $p = 0,111$ y PABAK = 0,128; $p = 0,107$) (7).

De hecho, en el ámbito de la TNE se necesita realizar estudios que permitan evaluar las prácticas asistenciales que afectan a la seguridad de los pacientes. Así pues, el objetivo de este estudio fue describir la evaluación de la concordancia entre observadores en la aplicación de una lista de control sobre la atención de los pacientes que utilizan TNE.

MÉTODO

Se trata de un estudio de fiabilidad realizado en 2018 en cuatro unidades de hospitalización de adultos de un hospital

universitario del sur de Brasil. Este estudio es una etapa que precedió a la realización de un ensayo clínico abierto, registrado en el *Clinical Trials* (NCT 03497221) y aprobado en cuanto a sus aspectos éticos y metodológicos en el marco de la CAAE: 63247916.5.0000.5327 (8).

El equipo de recogida de datos estaba formado por enfermeras, estudiantes universitarios y becarios de iniciación científica: en total, ocho asistentes de investigación (AI). Antes de la recolección de datos, una enfermera (una de las autoras, SMRS) llevó a cabo el entrenamiento del equipo con el objetivo de estandarizar: a) el acercamiento a los pacientes y la obtención del consentimiento para el estudio; b) la cumplimentación del instrumento de recogida de datos; c) la evaluación y el monitoreo de los resultados; y d) los registros en los formularios de búsqueda. Se utilizaron manuales para estandarizar todos los procedimientos de la investigación y los AI recibieron capacitaciones específicas para su inserción en el campo: A) buenas prácticas en la investigación clínica y seguridad del paciente en la investigación; B) buenas prácticas en TNE, cuando se presentó la lista de control (compuesta por 25 elementos de buenas prácticas en TNE). En esta última etapa: 1) se aclararon los objetivos del proyecto de investigación, la lista de control y las dudas acerca del vocabulario y/o de la interpretación de las preguntas; 2) se presentaron los materiales y los dispositivos utilizados en la gestión de la dieta enteral (bomba de infusión, jeringas, botellas de dieta); 3) revisión de la lista de control a la cabecera del paciente, como momento para aclarar dudas, con evaluación conjunta entre los AI y la enfermera de referencia, así como sobre los criterios de evaluación subjetiva de los datos de la lista de control. Después de la capacitación *in situ*, en la que los AI aplicaron la lista de control a la cabecera del paciente, supervisados directamente por la enfermera (SMRS), comenzó la evaluación de la concordancia entre observadores.

Cada par de observadores independientes, "cegados" para la evaluación opuesta, realizó las evaluaciones simultáneamente. Los registros de las evaluaciones se hicieron por medio de formularios electrónicos (Google Forms®), utilizando *tablets* o *smartphones*, y se enviaron a una base de datos a la que solo se tuvo acceso en el período de análisis de los datos.

La evaluación realizada por la enfermera (SMRS) se consideró el "estándar" ("evaluadora de referencia"), mientras que la evaluación realizada por el AI se consideró como "evaluación de prueba". Por cada evaluación de un AI había una evaluación estándar realizada por la enfermera (SMRS).

Para no influir en el análisis de la evaluadora de referencia, un miembro colaborador, ajeno a la investigación, ocultó los nombres de los AI en la base de datos, sustituyéndolos por un número (del 1 al 8). El cegamiento se mantuvo hasta el final del análisis de los datos.

La concordancia entre observadores se evaluó a partir del coeficiente kappa (κ) o mediante el coeficiente kappa ajustado (PABAK). El análisis de la concordancia de cada elemento de la lista de control y la concordancia global se realizó a través del programa *Statistical Package for the Social Sciences* – SPSS®, versión 20.0. Para calcular los intervalos de confianza (IC 95 %

se utilizó el programa *WinPepi*[®] o la calculadora *Single Case Research*[®], como se indica. Se utilizó la clasificación propuesta por Landis y Koch (3) para clasificar el nivel de concordancia con el valor de kappa: $k = 1,0$: de concordancia perfecta; $k = 0,99-0,81$: de concordancia casi perfecta; $k = 0,80-0,61$: de concordancia considerable; $k = 0,60-0,41$: de concordancia moderada; $k = 0,40-0,21$: de concordancia aceptable; $k = 0,2-0,1$: de concordancia ligera; y $k = 0,0$: de concordancia pobre (3).

Tras un análisis provisional de los datos, cuando hubo concordancias globales (de los elementos totales de la lista de control) o en elementos específicos de la lista de control con $k < 0,81$, los AI se sometieron a una nueva capacitación, así como a la repetición de todas las etapas de la evaluación por pares, lo que se denominó “momento 2” de un mismo AI. Los AI que no estaban disponibles para repetir la capacitación y la evaluación de la concordancia cuando estaba indicado, o que no alcanzaron una $k > 0,80$ después de la segunda ronda de capacitación, fueron considerados no aptos y, por lo tanto, no recogieron datos del ensayo clínico.

La concordancia se describió a nivel global (considerando el total de 25 elementos de la lista de control) y estratificada (para cada elemento de la lista de control). A efectos didácticos, la presentación de la concordancia de los elementos de control se agrupó por áreas temáticas: 1) los elementos relacionados con los procedimientos de infusión de la dieta y del agua por la SNE y con los cuidados de la bomba de infusión; 2) los elementos que se refieren a los materiales de apoyo en la gestión de la dieta enteral; 3) los elementos que se refieren al cuidado del paciente en el uso de la dieta enteral.

RESULTADOS

Se realizaron un total de 451 evaluaciones por pares entre la evaluadora de referencia (enfermera) y los ocho AI. Los AI 6 y 7 participaron en dos momentos diferentes de las evaluaciones, ya que no alcanzaron un nivel aceptable de concordancia en el primer bloque de evaluaciones. Así, los resultados de ambos AI se muestran en el momento 1 y en el momento 2.

El mayor número de evaluaciones se produjo con el AI 2 ($n = 54$); el menor, con el AI 7, en su segundo momento de

evaluación (momento 2) ($n = 28$). La concordancia general entre la enfermera y los AI fue casi perfecta en la mayoría de los pares de evaluaciones. Entre todos los pares de evaluaciones, la más concordante se dio con el AI 7 en el “momento 2” de la evaluación ($k = 0,96$; IC = 0,95-0,98) y la peor con el AI 8 ($k = 0,73$; IC = 0,70-0,77) (Tabla I).

Aunque la concordancia global con el AI 6 en el primer momento (momento 1) mostró una concordancia casi perfecta, en elementos específicos de la lista de control la concordancia fue menor (Tabla II). Esto requirió más entrenamiento y una segunda evaluación (momento 2), cuando se observaron mejoras en prácticamente todos los elementos relacionados con las botellas de dieta, el agua y la bomba de infusión para la gestión de la dieta enteral.

Se observa también que con la mayoría de los AI hubo un mayor número de concordancia ($kappa > 0,81$) en la evaluación de los elementos objetivos de la lista de control (fechas, validez y presencia o no de la botella de agua de hidratación). Por otra parte, la concordancia fue menor cuando se evaluaron los elementos subjetivos sobre la presencia de suciedad/residuos de cola en la bomba de infusión utilizada para la gestión de la dieta enteral (Tabla II).

La concordancia en la mayoría de las evaluaciones de los elementos relacionados con las condiciones de los materiales y dispositivos (presencia de suciedad, identificación y validez de las líneas de infusión, jeringas dietéticas y vasos desechables) utilizados para la gestión de la dieta se mostró casi perfecta o considerable. La excepción fueron algunos elementos (Tabla III) que sugieren una mala interpretación y que requerirían una recapacitación. Aun así, hubo un menor número de concordancias con el AI 7 en el primer momento de la evaluación (momento 1) en cuanto a elementos específicos de la lista de control. En un segundo momento (momento 2), después de una nueva capacitación, hubo una mejora de la concordancia con el AI 7 en casi todos los elementos relacionados con los materiales de apoyo en la gestión de la dieta (Tabla III).

Al evaluarse los elementos relacionados con el cuidado directo del paciente (cuidado de la posición de la cabeza y las condiciones de fijación de la SNE), la concordancia fue casi perfecta ($k > 0,81$), mostrándose considerable ($k = 0,61-0,80$) solo en la fecha de fijación de la sonda.

Tabla I. Concordancia global del total de los elementos de la lista de control entre la enfermera (evaluadora de referencia) y los asistentes de investigación (AI). Datos expresados por el valor del coeficiente kappa (k) y su intervalo de confianza (IC) del 95 %

| AI y número de observaciones | AI 1 n = 40 | AI 2 n = 54 | AI 3 n = 49 | AI 4 n = 45 | AI 5 n = 41 | AI 6 | | AI 7 | | AI 8 n = 48 |
|------------------------------|----------------|----------------|----------------|----------------|----------------|--------------------|--------------------|--------------------|--------------------|----------------|
| | | | | | | Momento 1 (n = 30) | Momento 2 (n = 34) | Momento 1 (n = 37) | Momento 2 (n = 28) | |
| Kappa (K) | 0,91 | 0,83 | 0,92 | 0,83 | 0,94 | 0,81 | 0,94 | 0,77 | 0,96 | 0,73 |
| IC 95 % | 0,89-0,93 | 0,80-0,85 | 0,90-0,94 | 0,80-0,86 | 0,92-0,96 | 0,78-0,85 | 0,92-0,96 | 0,73-0,80 | 0,95-0,98 | 0,70-0,77 |

AI: asistente de investigación; IC: intervalo de confianza; k: coeficiente kappa; 1,0: concordancia perfecta; 0,99-0,81: concordancia casi perfecta; 0,80-0,61: concordancia considerable; 0,60-0,41: concordancia moderada; 0,40-0,21: concordancia aceptable; 0,2-0,1: concordancia ligera; 0,0: concordancia pobre.

Tabla II. Concordancia entre la enfermera (evaluadora de referencia) y los asistentes de investigación (AI) en los elementos de la lista de control referentes a las infusiones de dieta y de agua, y a la bomba de infusión utilizada en la terapia nutricional enteral (TNE)

| Elementos de la lista de control | AI 1 n = 40 | AI 2 n = 54 | AI 3 n = 49 | AI 4 n = 45 | AI 5 n = 41 | AI 6 | | | AI 7 | AI 8 |
|---|----------------|----------------|----------------|----------------|----------------|---------------------|---------------------|---------------------|---------------------|-----------|
| | | | | | | Momento 1 n = 30 | Momento 2 n = 34 | Momento 1 n = 37 | Momento 2 n = 28 | n = 48 |
| Botella de dieta identificada igual que el brazalete del paciente (k) | 1,0 | 0,87 | 0,81 | 1,0 | 1,0 | 0,71 | 1,0 | 0,84 | 1,0 | 1,0 |
| IC 95 % | 0,85-1,0* | 0,69-1,0* | 0,6-1,0 | 0,86-1,0* | 0,86-1,0* | 0,35-1,0 | 0,84-1,0* | 0,54-1,0 | 0,83-1,0* | 0,87-1,0* |
| Botella de dieta gestionada dentro de la validez (3 h) (k) | 0,61 | 0,58 | 0,96 | 0,35 | 0,95 | 0,71 | 1,0 | 0,49 | 0,93 | 0,73 |
| IC 95 % | 0,52-0,81 | 0,54-0,79 | 0,84-1,0 | 0,43-0,7 | 0,81-1,0 | 0,63-0,97 | 0,84-1,0* | 0,40-0,70 | 0,81-1,0 | 0,68-0,95 |
| Botella de agua para lavar la SNE/línea de infusión igual que el brazalete del paciente (k) | 0,91 | 0,96 | 0,96 | 0,79 | 1,0 | 0,92 | 1,0 | 0,95 | 1,0 | 0,58 |
| IC 95 % | 0,73-1,0 | 0,76-1,0 | 0,87-1,0 | 0,6-0,99 | 0,86-1,0* | 0,76-1,0 | 0,84-1,0* | 0,83-1,0 | 0,83-1,0* | 0,39-0,76 |
| Botella de agua para lavar la SNE/ línea de infusión en la validez (24 h) (k) | 0,95 | 0,89 | 0,97 | 0,68 | 0,92 | 0,89 | 1,0 | 0,92 | 1,0 | 0,71 |
| IC 95 % | 0,85-1,0 | 0,76-1,0 | 0,84-1,0 | 0,56-0,84 | 0,78-1,0 | 0,73-1,0 | 0,84-1,0* | 0,77-1,0 | 0,83-1,0* | 0,56-0,87 |
| Botella de agua para hidratación igual que el brazalete del paciente (k) | 0,95 | 0,93 | 0,92 | 0,96 | 0,95 | 0,86 | 0,94 | 0,79 | 1,0 | 0,50 |
| IC 95 % | 0,86-1,0 | 0,82-1,0 | 0,81-1,0 | 0,87-1,0 | 0,85-1,0 | 0,67-1,0 | 0,82-1,0 | 0,69-0,99 | 0,83-1,0* | 0,37-0,63 |
| Botella del agua de hidratación dentro de la validez (24 h) (k) | 0,87 | 0,93 | 0,93 | 0,88 | 0,95 | 0,86 | 0,94 | 0,69 | 1,0 | 0,42 |
| IC 95 % | 0,74-1,0 | 0,82-1,0 | 0,81-1,0 | 0,76-1,0 | 0,82-1,0 | 0,67-1,0 | 0,82-1,0 | 0,56-0,87 | 0,83-1,0* | 0,28-0,57 |
| Bomba de infusión limpia y sin residuos (k) | 1,0 | 0,62 | 0,95 | 0,75 | 1,0 | 0,61 | 0,88 | 0,84 | 1,0 | 0,86 |
| IC 95 % | 0,85-1,0* | 0,62-0,87 | 0,85-1,0 | 0,69-0,97 | 0,86-1,0* | 0,58-0,92 | 0,71-1,0 | 0,73-1,0 | 0,83-1,0* | 0,71-1,0 |

AI: asistente de investigación; IC: intervalo de confianza; k: coeficiente kappa; 1,0: concordancia perfecta; 0,99-0,81: concordancia casi perfecta; 0,80-0,61: concordancia considerable; 0,60-0,41: concordancia moderada; 0,40-0,21: concordancia aceptable; 0,2-0,1: concordancia ligera; 0,0: concordancia pobre. *Los valores del intervalo de confianza matemáticamente calculados como > 1 se consideran = 1.

Tabla III. Concordancia entre la enfermera (evaluadora de referencia) y los asistentes de investigación (AI) en los elementos de la lista de control referentes a los materiales de apoyo para la gestión de la terapia nutricional enteral (TNE)

| Elementos de la lista de control | AI 1 n = 40 | AI 2 n = 5 | AI 3 n = 49 | AI 4 n = 45 | AI 5 n = 41 | AI 6 | | AI 7 | | AI 8 n = 48 |
|---|----------------|---------------|----------------|----------------|----------------|---------------------|---------------------|---------------------|---------------------|----------------|
| | | | | | | Momento 1 n = 30 | Momento 2 n = 34 | Momento 1 n = 37 | Momento 2 n = 28 | |
| Línea de infusión azul o púrpura (k) | 1,0 | 0,80 | 0,91 | 0,81 | 1,0 | 0,65 | 1,0 | 1,0 | 1,0 | 1,0 |
| IC (95 %) | 0,85-1,0* | 0,6-1,0 | 0,74-1,0 | 0,55-1,0 | 0,86-1,0* | 0,02-1,0 | 0,84-1,0* | 0,85-1,0* | 0,83-1,0* | 0,87-1,0* |
| Línea de infusión dentro de la validez (24 h) (k) | 1,0 | 0,87 | 0,96 | 0,76 | 1,0 | 0,74 | 0,93 | 0,76 | 0,93 | 0,96 |
| IC (95 %) | 0,85-1,0* | 0,76-1,0 | 0,84-1,0 | 0,66-0,94 | 0,86-1,0* | 0,68-1,0 | 0,80-1,0 | 0,65-0,95 | 0,79-1,0 | 0,84-1,0 |
| Línea de infusión limpia o sin residuos cuando se pausa la dieta (k) | 0,92 | 0,08 | 0,88 | 0,33 | 0,62 | 0,50 | 0,80 | 0,62 | 1,0 | 0,53 |
| IC (95 %) | 0,81-1,0 | 0,15-0,40 | 0,81-1,0 | 0,50-0,77 | 0,71-0,99 | 0,33-0,92 | 0,53-1,0 | 0,52-0,83 | 0,83-1,0* | 0,33-0,74 |
| Punta de la línea de infusión protegida con tapa cuando se pausa la dieta (k) | 0,81 | 0,74 | 0,91 | 0,54 | 0,78 | 0,63 | 1,0 | 0,47 | 1,0 | 0,75 |
| IC (95 %) | 0,63-0,98 | 0,46-1,0 | 0,74-1,0 | 0,07-1,0 | 0,78-1,0 | 0,58-0,92 | 0,84-1,0* | 0,36-0,66 | 0,83-1,0* | 0,55-0,96 |
| Jeringa para TNE (Oralpak®) en la cama del paciente (k) | 1,0 | 0,87 | 0,97 | 1,0 | 1,0 | 0,83 | 1,0 | 0,92 | 1,0 | 0,45 |
| IC (95 %) | 0,85-1,0* | 0,76-0,99 | 0,84-1,0 | 0,86-1,0* | 0,86-1,0* | 0,68-1,0* | 0,84-1,0* | 0,77-1,0 | 0,83-1,0* | 0,31-0,58 |
| Jeringa Oralpak® con el nombre/prontuario del paciente (k) | 0,96 | 0,87 | 0,96 | 0,96 | 1,0 | 0,80 | 1,0 | 0,78 | 1,0 | 0,52 |
| IC (95 %) | 0,89-1,0 | 0,75-0,99 | 0,84-1,0 | 0,83-1,0 | 0,86-1,0* | 0,63-0,97 | 0,84-1,0* | 0,65-0,95 | 0,83-1,0* | 0,31-0,72 |
| Jeringa Oralpak® etiquetada con fecha/turno del paciente (k) | 0,92 | 0,83 | 0,96 | 0,96 | 1,0 | 0,80 | 1,0 | 0,72 | 0,87 | 0,52 |
| IC (95 %) | 0,81-1,0 | 0,73-0,98 | 0,84-1,0 | 0,83-1,0 | 0,86-1,0* | 0,63-0,97 | 0,84-1,0* | 0,60-0,91 | 0,70-1,0 | 0,31-0,72 |
| Jeringa Oralpak® dentro de la validez (24 h) (k) | 0,96 | 0,86 | 0,89 | 0,90 | 1,0 | 0,80 | 0,95 | 0,72 | 0,94 | 0,52 |
| IC (95 %) | 0,88-1,0 | 0,76-1,0 | 0,78-1,0 | 0,76-1,0 | 0,86-1,0* | 0,63-0,97 | 0,54-0,82 | 0,60-0,91 | 0,82-1,0 | 0,33-0,73 |
| Jeringa Oralpak® limpia y sin residuo (k) | 0,92 | 0,67 | 0,9 | 0,80 | 0,92 | 0,60 | 0,82 | 0,63 | 0,94 | 0,52 |
| IC (95 %) | 0,82-1,0 | 0,54-0,79 | 0,81-1,0 | 0,70-0,97 | 0,78-1,0 | 0,43-0,77 | 0,66-0,98 | 0,48-0,79 | 0,82-1,0 | 0,34-0,70 |
| Vaso de plástico para la higienización limpio y sin residuos (k) | 0,87 | 0,75 | 0,6 | 1,0 | 0,81 | 0,86 | 1,0 | 0,46 | 1,0 | 0,82 |
| IC (95 %) | 0,74-1,0 | 0,62-0,87 | 0,56-0,83 | 0,86-1,0* | 0,67-0,96 | 0,68-1,0 | 0,84-1,0* | 0,28-0,58 | 0,83-1,0* | 0,68-0,97 |
| Vaso de plástico seco (k) | 0,90 | 0,9 | 0,92 | 0,96 | 0,95 | 0,61 | 1,0 | 0,90 | 1,0 | 0,84 |
| IC (95 %) | 0,77-1,0 | 0,79-1,0 | 0,81-1,0 | 0,83-1,0 | 0,82-1,0 | 0,48-0,82 | 0,84-1,0* | 0,77-1,0 | 0,83-1,0* | 0,69-0,98 |
| Vaso de plástico dentro de la validez de 24 h (k) | 0,82 | 0,82 | 0,92 | 0,96 | 0,90 | 0,94 | 1,0 | 0,95 | 1,0 | 0,78 |
| IC (95 %) | 0,66-0,98 | 0,66-0,97 | 0,81-1,0 | 0,87-1,0 | 0,77-1,0 | 0,78-1,0 | 0,84-1,0* | 0,81-1,0 | 0,83-1,0* | 0,59-0,96 |

AI: asistente de investigación; IC: intervalo de confianza; k: coeficiente kappa; 1,0: concordancia perfecta; 0,99-0,81: concordancia casi perfecta; 0,80-0,61: concordancia considerable; 0,60-0,41: concordancia moderada; 0,40-0,21: concordancia aceptable; 0,2-0,1: concordancia ligera; 0,0: concordancia pobre.

*Los valores del intervalo de confianza matemáticamente calculados como > 1 se consideran = 1.

Tabla IV. Concordancia entre la enfermera (evaluadora de referencia) y los asistentes de investigación (AI) en los elementos de la lista de control relativos al cuidado del paciente junto a la cama con terapia nutricional enteral (TNE)

| Elementos de la lista de control | AI 1 n = 40 | AI 2 n = 54 | AI 3 n = 49 | AI 4 n = 45 | AI 5 n = 41 | AI 6 | | AI 7 | | AI 8 n = 48 |
|---|----------------|----------------|----------------|----------------|----------------|---------------------|---------------------|---------------------|---------------------|----------------|
| | | | | | | Momento 1 n = 30 | Momento 2 n = 34 | Momento 1 n = 37 | Momento 2 n = 28 | |
| Fijación limpia de la SNE (k) | 0,95 | 0,53 | 0,60 | 0,32 | 0,75 | 0,66 | 0,70 | 0,66 | 1,0 | 0,41 |
| IC (95 %) | 0,85-1,0 | 0,36-0,75 | 0,4-0,8 | 0,01-0,63 | 0,55-0,96 | 0,36-0,96 | 0,46-0,94 | 0,42-0,91 | 0,83-1,0* | 0,16-0,67 |
| Fijación sin oleosidad de la SNE (k) | 0,62 | 0,51 | 0,70 | 0,66 | 0,80 | 0,70 | 0,36 | 0,21 | 0,92 | 0,37 |
| IC (95 %) | 0,29-0,95 | 0,25-0,78 | 0,5-0,94 | 0,42-0,91 | 0,57-1,0 | 0,43-0,97 | 0,01-0,71 | -0,11-0,53 | 0,78-1,0 | 0,06-0,68 |
| Fijación de la SNE datada en hasta 24 h anteriores (k) | 0,84 | 0,92 | 0,90 | 0,74 | 0,73 | 0,76 | 0,93 | 0,66 | 0,93 | 0,67 |
| IC (95 %) | 0,55-1,0 | 0,77-1,0 | 0,78-1,0 | 0,51-0,98 | 0,37-1,0 | 0,50-1,0 | 0,79-1,0 | 0,65-0,95 | 0,80-1,0 | 0,37-0,97 |
| Fijación de la SNE no despegada (k) | 0,54 | 0,70 | 0,82 | 0,49 | 0,80 | 0,63 | 0,40 | 0,72 | 0,65 | 0,67 |
| IC (95 %) | 0,22-0,85 | 0,51-0,89 | 0,7-0,99 | 0,23-0,75 | 0,62-0,99 | 0,34-0,92 | 0,05-0,74 | 0,49-0,95 | 0,40-0,91 | 0,46-0,87 |
| Fijación de la SNE no tirada (k) | 1,0 | 0,44 | 1,0 | 1,0 | 0,36 | 0,65 | 0,80 | -0,07 | 0,89 | 0,03 |
| IC (95 %) | 0,85-1,0* | 0,06-0,81 | 0,87-1,0 | 0,86-1,0* | 0,02-0,77 | 0,02-1,0 | 0,53-1,0 | -0,14-0,00 | 0,68-1,0 | -0,07-0,01 |
| La cama del paciente a $\geq 30^\circ$ durante la infusión de la dieta/agua (k) | 0,57 | 0,74 | 0,81 | 0,59 | 0,89 | 0,31 | 0,82 | 0,48 | 0,87 | 0,33 |
| IC (95 %) | 0,52-0,81 | 0,55-0,93 | 0,64-0,99 | 0,32-0,87 | 0,75-1,0 | 0,53-0,87 | 0,58-1,0 | 0,40-0,71 | 0,71-1,0 | 0,49-0,76 |

AI: asistente de investigación; IC: intervalo de confianza; k: coeficiente kappa; 1,0: concordancia perfecta; 0,99-0,81: concordancia casi perfecta; 0,80-0,61: concordancia considerable; 0,60-0,41: concordancia moderada; 0,40-0,21: concordancia aceptable; 0,2-0,1: concordancia ligera; 0,0: concordancia pobre.

*Los valores del intervalo de confianza matemáticamente calculados como > 1 se consideran = 1.

Al evaluarse los elementos más subjetivos, como la presencia de suciedad en la fijación de la SNE o de oleosidad y la fijación de la SNE desprendida o tirada, la comparación en la mayoría de los elementos fue débil, relativa o moderada. Además, "la posición de la cama del paciente (igual o $> 30^\circ$ ") mostró una concordancia de moderada a aceptable ($k < 0,49$) en tres de los ocho observadores (Tabla IV).

En vista de los resultados obtenidos, considerando que la concordancia entre la enfermera y los AI 1, 3 y 5 fue casi perfecta ($k > 0,81$) para la mayoría de los elementos de la lista de control, esos AI realizaron la recopilación de datos en el ensayo clínico que sucedió al presente estudio. La concordancia con los AI 2, 4, 6, 7 y 8 fue menor ($k < 0,81$) en muchos elementos específicos

de la lista de control, lo que requirió una nueva capacitación y una nueva evaluación de la concordancia. Los AI 6 y 7, después de una nueva capacitación y evaluación, mostraron una concordancia casi perfecta ($k > 0,81$) y participaron en la recogida de datos, mientras que los AI 2 y 4 no continuaron en el proyecto y el AI 8 se capacitó para actividades distintas de la aplicación de la lista de control durante el ensayo clínico.

DISCUSIÓN

El presente estudio mostró una excelente concordancia entre observadores en la aplicación de una lista de control para la aten-

ción de los pacientes que utilizan una sonda nasointestinal, cuando se les capacitó y supervisó para reunir datos de una investigación clínica. Este estudio señaló también que la evaluación de la concordancia resultó fundamental para garantizar la fiabilidad de la recogida de datos para el ensayo clínico realizado a continuación, ya que algunos asistentes de investigación requirieron capacitación y otro mostró una concordancia insuficiente. De esa manera, si no se hubiera hecho la evaluación de la concordancia, el ensayo clínico podría haberse visto afectado por un error sistemático de contraste.

Además, el estudio documentó la necesidad de evaluar detalladamente la concordancia entre evaluadores a fin de identificar los elementos cuya observación por parte de los asistentes de investigación puede verse comprometida por criterios de evaluación subjetivos, en los que existe una mayor posibilidad de discrepancias, aunque igualmente importante en cuanto al rigor metodológico.

Se obtuvieron resultados similares a los nuestros en un estudio realizado en China (9). Los autores evaluaron la concordancia entre una enfermera y un radiólogo al interpretar una imagen de rayos X para detectar la posición anatómica del extremo distal de un catéter central. Ambos evaluaron 599 imágenes y la concordancia entre observadores se consideró excelente ($k = 0,93$; $p < 0,001$; IC 95 %: 0,90-0,97). En otro estudio (10), en el que se evaluó la concordancia entre nueve patólogos para identificar la presencia de linfocitos infiltrados de tumores de mama en 124 portaobjetos, se encontró una concordancia baja (k entre 0,38 y 0,46). Frente a esos resultados, los autores llegaron a la conclusión de que esa prueba diagnóstica no debería aplicarse en la práctica asistencial (10). Los dos estudios presentados muestran los extremos de la concordancia entre observadores y cómo esta evaluación interfiere en la toma de decisiones sobre una práctica asistencial. Lo mismo puede extrapolarse al campo de la investigación, ya que la baja fiabilidad de los datos generados por un AI o más puede sesgar el estudio, comprometiendo la totalidad de los resultados.

Nuestro estudio demostró que la capacitación se mostró fundamental para obtener mayores puntuaciones de concordancia entre observadores en el momento 2. En otro estudio que comparó la concordancia entre observadores radiólogos con experiencia limitada en la evaluación de imágenes de resonancia magnética abdominal se compararon con expertos en la interpretación de imágenes de resonancia magnética en pacientes con sospecha de apendicitis (11). En ese estudio hubo una formación previa de los no expertos que se asemeja al método empleado por nosotros (capacitación previa de los AI); aunque la concordancia obtenida fue considerable ($k = 0,78$), se mostró inferior a la del presente estudio. Eso demuestra que, en estudios de esta magnitud, cuando la concordancia entre observadores es baja, se debe considerar la evaluación de la necesidad de capacitación.

Una posible limitación de nuestro estudio es que el rol del AI fuera desempeñado por estudiantes universitarios de enfermería y no todos son del mismo período de formación. Eso puede haber contribuido, en parte, a que algunos AI hayan obtenido resultados discrepantes en algunos de los elementos de la lista de control,

aunque todos hayan sido sometidos a una capacitación presencial y a una supervisión directa, hasta que se dispusiera de conocimientos y aptitudes específicas para aplicar la lista de control. Por otra parte, el rigor metodológico adoptado por nosotros para insertar los AI en el campo de estudio, la creación de manuales de procedimientos de investigación, su capacitación y la supervisión directa fueron capaces de garantizar la calidad de la recogida de datos en el ensayo clínico que siguió (12).

CONCLUSIONES

La evaluación de la concordancia entre observadores es esencial para garantizar la fiabilidad de los datos recogidos en los estudios clínicos epidemiológicos. La evaluación de la fiabilidad entre observadores es una de las medidas importantes para reducir al mínimo el riesgo de sesgo en los estudios, lo que aumenta la solidez de las pruebas generadas mediante estudios metodológicos sólidos.

APROBACIÓN DEL COMITÉ DE ÉTICA DE LA INVESTIGACIÓN

Aprobado por el Comité de Ética de investigación del Hospital de Clínicas de Porto Alegre (vinculado a la Universidad Federal de Rio Grande do Sul — UFRGS), con el número 16-0534, en la plataforma Brasil bajo CAAE 63247916500005327, y registrado en Clinical Trials (NCT 03497221).

AGRADECIMIENTOS

A nuestro grupo de investigación, Nursing Care Study Group (NUCAS), del programa de posgrado en enfermería de la Universidad Federal de Rio Grande do Sul (PPGENF/UFRGS); al Grupo de Investigación y Posgrado del Hospital de Clínicas de Porto Alegre (GPPG/HCPA) para consultorías estadísticas, presupuestarias y de presentación a CONEP/CEP; a la Coordinación para la Mejora del Personal de Educación Superior (CAPES).

BIBLIOGRAFÍA

1. Agencia Nacional de Vigilância Sanitária. RCD N° 63, de 6 de julho de 2000. ANVISA; 2000.
2. Alpendre FT, Cruz ED de A, Dyniewicz AM, Mantovani M de F, Silva AEB de C, Santos G de S dos. Safe surgery: validation of pre and postoperative checklists. *Rev Lat Am Enfermagem* 2017;25:e2907. DOI: 10.1590/1518-8345.1854.2907
3. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74. DOI: 10.2307/2529310
4. Lima EMB, Almeida BL, Gomes HB, Bartochevis JAB, Toniato TS, Lazzarini TR, et al. Agreement between Graz Malnutrition Screening (GMS) with subjective nutritional assessment instruments in hospitalized patients. *Nutr Hosp* 2018;35:1138-44. DOI: 10.20960/nh.1853

5. Cerqueira CT de, Mello MJG de, Viana L de A, Macêdo DJ do N, Figueiroa JN. Comparison of weight estimation methods in hospitalized Brazilian children and adolescents. *Nutr Hosp* 2020;37:243-50. DOI: 10.20960/nh.02594
6. Azambuja FB, Beghetto MG, de Assis MCS, de Mello ED. Food intake reported versus nursing records: is there agreement in surgical patients? *Nutr Hosp* 2015;31:2735-9. DOI: 10.3305/nh.2015.31.6.8691
7. Beghetto MG, Anziliero F, Leães DM, Mello ED de. Sondagem enteral: concordância entre teste de ausculta e raio-x na determinação do posicionamento da sonda. *Rev Gaúcha Enferm* 2015;36:98-103. DOI: 10.1590/1983-1447.2015.04.54700
8. Silva SMR. Impacto de uma campanha de identidade visual sobre o processo de administração de dieta por sonda nasoenteral e sobre a segurança do paciente: ensaio clínico aberto [Tese de Doutorado]. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2019.
9. Zhang X, Jia D, Ke N, Liu C, Fu L, Hu X. Excellent inter-observer agreement between radiologist and nurse: tracheal carina-based identification of peripherally inserted central catheter tip position. *J Vasc Access* 2018;19:28-33. DOI: 10.5301/jva.5000799
10. Tramm T, Di Caterino T, Jylling A-MB, Lelkaitis G, Lænkholm A-V, Ragó P, et al. Standardized assessment of tumor-infiltrating lymphocytes in breast cancer: an evaluation of inter-observer agreement between pathologists. *Acta Oncol* 2018;57:90-4. DOI: 10.1080/0284186X.2017.1403040
11. Leeuwenburgh MMN, Wiarda BM, Jensch S, Van Es HW, Stockmann HBAC, Willem J, et al. Accuracy and interobserver agreement between MR-non-expert radiologists and MR-experts in reading MRI for suspected appendicitis. *Eur J Radiol* 2014;83:103-10. DOI: 10.1016/j.ejrad.2013.09.022
12. World Health Organization. Calibration of examiners for oral health epidemiological surveys. Geneva: ORH/EPID; 1993.