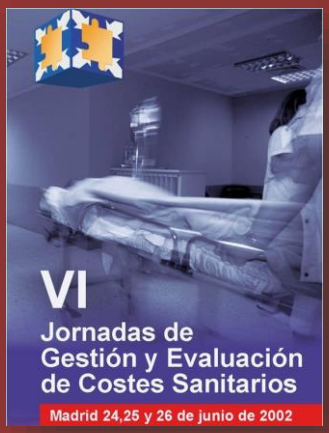

COMUNICACIONES

Título

FIABILIDAD DE UN SISTEMA PARA EVALUAR LOS RESULTADOS DE LOS SERVICIOS DE UN HOSPITAL

Clasificación

08 PLANIFICACION Y EMPRESA / 09 EMPRESA

Forma de presentación: COMUNICACIÓN

Palabras clave: EVALUACIÓN DE SERVICIOS; VALIDACIÓN; DIRECCIÓN POR OBJETIVOS

Autores

BARRASA VILLAR, JI., ARNAL SORIANO, JA., MOLINER LAHOZ, FJ., PRADOS TORRES, A., CALLEN SEVILLA, L., RAMIREZ GASCA, T., GARCIA MONTERO, JI., SANCHEZ ARROYO, R.

Entidad HOSPITAL CLINICO UNIVERSITARIO "LOZANO BLESA"

CCAA ARAGON

INTRODUCCION

Hoy por hoy parece un hecho mayoritariamente aceptado por gestores, políticos e incluso profesionales, que es necesario implantar en el sistema sanitario español mecanismos de evaluación que sirvan para arbitrar la distribución de incentivos entre los diferentes equipos de trabajo, vinculando su actividad con el cumplimiento de determinados objetivos institucionales.

La introducción de sistemas de retribución variable, el desarrollo de la carrera profesional, o la concesión de premios y méritos necesita de esos mecanismos de evaluación para su perfecto encaje y desarrollo.

Sin embargo, la elaboración de este tipo de herramientas es muy complicada, tanto desde el punto de vista técnico como desde la más delicada perspectiva de las relaciones laborales.

La mayor parte de estas dificultades están relacionadas con la mayor o menor fiabilidad de los resultados obtenidos. Cualquier sistema de evaluación que se quiera implantar debe ofrecer confianza tanto a evaluadores como a evaluados. De lo contrario se convertirá en una fuente de conflicto laboral de efectos imprevisibles y contraproducentes.

Además de la puesta en marcha de mecanismos de participación que persigan el consenso con los servicios evaluados del procedimiento de evaluación a seguir, una buena forma de garantizar esta confianza es mediante el desarrollo de herramientas que sean fiables y objetivas desde el punto de vista del método e indicadores empleados, como de los resultados que con ella vamos a obtener.

En este sentido, el objetivo que nos planteamos en este trabajo era someter el instrumento de evaluación utilizado en nuestro hospital para el reparto de la productividad variable del año

2001 a un proceso de validación en el que se hiciera explícito su grado de fiabilidad examinando los dos determinantes principales de esa cualidad en cualquier herramienta de medición: la precisión y la exactitud.

MATERIAL Y METODOS

La duración del proyecto ha sido de 20 meses, desde septiembre de 2000 hasta mediados de abril de 2002. En las primeras fases se desarrolló la herramienta de evaluación; para ello, se definieron los atributos del sistema de forma explícita (representatividad, pertinencia, claridad, transparencia, accesibilidad, veracidad, estabilidad, carácter reproducible y eficiencia) y se utilizaron dichos criterios para seleccionar indicadores y desarrollar los modelos de evaluación específicos de cada servicio. Posteriormente se procedió a la aplicación y validación del método y a la difusión de sus resultados.

Para mejorar la precisión del sistema se hicieron definiciones operativas de todos los indicadores, se automatizó el proceso de cálculo y se extrajo la información de los cuadros de mando habituales del hospital, de tal forma que quienes se encargaban de su recopilación y elaboración eran personas muy entrenadas y acostumbradas a realizar esa tarea y desconocían por completo el uso paralelo que se iba a dar a la información que recogían. Además, una misma persona completó la evaluación en dos ocasiones diferentes para examinar la coherencia intraobservador. Una vez terminada la evaluación, los resultados individuales, servicio a servicio, fueron presentados al equipo directivo y a los jefes de servicio, que disponían todos ellos de la información necesaria para poder recalcular y reproducir la evaluación. Las discrepancias que pudieron darse se utilizaron como reflejo de la coherencia interobservador. Para tener una idea general de la precisión del modelo se calculó también el coeficiente de variación (CV) con las puntuaciones totales y por apartados (calidad, costes, actividad) que consiguieron los servicios y su bondad de ajuste a una distribución normal.

La exactitud (validez de construcción y de criterio) del sistema se ha fundamentado sobre todo en el cumplimiento de los criterios de validación de los indicadores y en la negociación y consenso de los tramos de variación aceptables en ellos con los servicios. No obstante, antes de empezar el diseño del sistema, se solicitó a la Comisión de Dirección del centro que hiciera de forma individual y anónima una evaluación de los diferentes servicios calificando a cada uno de ellos de acuerdo con su propia experiencia directiva. Al no disponer de ningún gold standard utilizamos esa opinión como patrón de comparación y examinamos la correlación existente entre las dos evaluaciones, así como su grado de concordancia (indice kappa) a la hora de clasificar a los servicios entre los 10 primeros y entre los 10 últimos. Además, en el proceso de análisis estadístico utilizamos la prueba de Kolmogorov-Smirnov para comprobar la hipótesis de ajuste a una distribución normal, los test de Kruskal-Wallis y de Wilcoxon para contrastar las diferencias de medias en muestras independientes y apareadas, respectivamente y el coeficiente de Pearson para demostrar la existencia de correlación entre variables cuantitativas.

En todos los casos, el máximo nivel de error aleatorio que admitimos fue del 5 %.

CONCLUSIONES

Se evaluaron un total de 56 unidades asistenciales, cuya calificación media fue de 73,5±18,1 puntos (sobre 100), con un mínimo de 29 y un máximo de 100 puntos. Los servicios Centrales (82,9±10,1) tuvieron una mejor calificación que los Quirúrgicos (70,5±18.5) y éstos que los Médicos (68,7±20,1).

En Actividad la puntuación media fue de 52,1±14,7 sobre 70, en Costes de 13,7±7,4 sobre 20 y en Calidad de 7,6±2,1 sobre 10. El modelo se ajustaba a una distribución normal, con un CV de 0,25, un margen de error intraobsevador de 0,25% e interobservador del 0,48 %. Por su parte la correlación entre la opinión del equipo directivo (estándar de comparación) y los resultados de la evaluación era buena y significativa (r= 0,43 p=0,004), aunque los índices kappa para medir la concordancia en la clasificación de los 10 primeros y 10 últimos puestos era baja (0,21).

En resumen, el sistema de evaluación contrastado tenía una buena precisión y capacidad de discriminación como lo demuestran el escaso margen de error intra e interobservador, los bajos coeficientes de variación de las puntuaciones y el buen ajuste a una distribución normal.

La validez de construcción del sistema estaba garantizada por el proceso de validación de los indicadores, en el que se aseguraba de forma apropiada y explícita la representación de todas las dimensiones relevantes de cada servicio en los apartados de asistencia, calidad y coste.

Con respecto a la validez de criterio, aunque la opinión de la Comisión de Dirección puede parecer un "gold standard" muy subjetivo, en realidad es el mejor y quizás el único que podemos disponer. El sistema probado tenía una buena correlación con él en cuanto a los puntuaciones otorgadas a los servicios, pero escasa concordancia en cuanto al establecimiento de los primeros y últimos puestos clasificatorios.

En conjunto puede concluirse que el sistema de evaluación probado es preciso, está bien construido y parece que su validez de criterio es bastante aceptable, lo que le confiere un grado de fiabilidad suficiente como para seguir usándolo y mejorándolo en años sucesivos.