

Copyright © 2015 Banco Interamericano de Desarrollo. Esta obra está bajo una licencia Creative Commons IGO 3.0 Reconocimiento-No Comercial-Sin Obra Derivada (CC-IGO BY-NC-ND 3.0 IGO) (<http://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) y puede ser reproducida para cualquier uso no-comercial otorgando crédito al BID.  No se permiten obras derivadas.

Cualquier disputa relacionada con el uso de las obras del BID que no pueda resolverse amistosamente se someterá a arbitraje de conformidad con las reglas de la CNUDMI. El uso del nombre del BID para cualquier fin que no sea para la atribución y el uso del logotipo del BID, estará sujeta a un acuerdo de licencia por separado y no está autorizado como parte de esta licencia CC-IGO.

Notar que el enlace URL incluye términos y condicionales adicionales de esta licencia.

Las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del Banco Interamericano de Desarrollo, de su Directorio Ejecutivo ni de los países que representa.

**Script: Aspectos técnicos de la recolección de datos**

**Diapositiva 1**

Hola. Me llamo Patricia Yañez-Pagans y en este video te voy a explicar algunos aspectos técnicos que se deben tener en cuenta a la hora de recolectar datos.

**Diapositiva 2**

El proceso de recolección de datos consiste en la recopilación de información necesaria para cuantificar los impactos de una intervención. Existen dos tipos de datos. Los datos de fuente primaria, que son aquellos recolectados de primera mano, como por ejemplo una encuesta de beneficiarios y no beneficiarios levantada por el equipo de evaluación. Los datos de fuente secundaria, son datos ya existentes o levantados por fuentes externas al equipo de evaluación, como por ejemplo censos de población, datos administrativos, o encuestas de hogares a nivel nacional.

En general, los datos primarios y secundarios son complementarios en una evaluación de impacto. De todas formas, consideraciones de costos así como otras características de los datos existentes, entre ellas el tamaño de la muestra, la cobertura y frecuencia de los mismos pueden ayudar a decidir si estos datos son suficientes para la evaluación.

**Diapositiva 3**

Es útil pensar en la cadena de resultados para saber qué tipo de datos recolectar. El primer paso consiste en recolectar información sobre la intervención. Por ejemplo, para evaluar los impactos de un programa de tutorías, será necesario saber qué alumnos reciben o no tutorías, cuántas horas de tutoría recibió cada alumno o qué alumnos recibieron las tutorías primero.

El siguiente paso es identificar los resultados de la intervención que pueden ser directos o indirectos. Es importante distinguir entre efectos de corto plazo, como ser las calificaciones de los alumnos en pruebas estandarizadas, o efectos de largo plazo como ser el índice de permanencia en la escuela o el índice de inserción laboral.

Finalmente, será importante recolectar información sobre otros factores exógenos que podrían afectar los resultados y que no están directamente afectados por la intervención, como ser el sexo, la edad o la educación de los padres.

**Diapositiva 4**

Cuando recolectamos información sobre indicadores de resultados y otras variables, debemos tener siempre presente que la información sea de buena calidad. Para esto, el criterio básico será que los indicadores sean SMART.

Esto significa que deberán ser *específicos*, por ejemplo, si las tutorías son en matemáticas las pruebas de aprendizaje deberán medir este conocimiento y no otro.

Los indicadores deben ser *medibles,* esto significa que debe existir un instrumento de medición que sea claro y homogéneo para todos, como por ejemplo las pruebas estandarizadas de aprendizaje.

Los indicadores deben ser *atribuibles*, lo cual significa que debemos recolectar información sobre variables que muy posiblemente se vean afectadas por la intervención. Por ejemplo, las tutorías muy posiblemente tengan impactos directos sobre calificaciones, pero no tengan efectos en la salud de los alumnos.

Los indicadores deben ser *realistas*, esto significa que la recolección de información sobre los mismos debe ser factible y simple.

Los indicadores deben ser *focalizados*. Esto significa que la información debe ser recolectada para los grupos y contextos adecuados. Por ejemplo, la información sobre pruebas de matemáticas debe ser recolectada para escuelas donde existen alumnos con y sin tutoría.

**Diapositiva 5**

La importancia de recolectar información sobre otros factores externos puede ilustrarse con la siguiente figura. Supongamos que sólo contamos con información sobre las calificaciones de los alumnos con y sin tutoría. Si la participación en tutorías es aleatoria, el impacto se mide comparando el promedio de calificaciones de los alumnos en el grupo T o con tutoría y los alumnos en el grupo C sin tutoría o grupo de control.

Asumamos ahora que recolectamos información sobre la edad de los alumnos y su sexo. Este tipo de información nos permite reagrupar a los alumnos en subgrupos más homogéneos y medir los impactos dentro de cada grupo. Por ejemplo, el impacto de los niños en la región 1 y que son mayores será la comparación de calificaciones dentro de este subgrupo.

El impacto final se calcula como una media ponderada de la diferencia de calificaciones entre grupos de tratamiento y de control dentro de cada subgrupo. Este cálculo reduce la variación en las estimaciones de impacto haciendo que sean más precisas.

**Diapositiva 6**

Una vez que tenemos claro qué tipo de información necesitamos recolectar, el siguiente paso es definir los tiempos de recolección. Si el presupuesto lo permite, es siempre recomendable levantar una *encuesta de línea de base* antes de implementar la intervención. Esto es útil para recolectar datos que no han sido afectados por la intervención y permite evaluar qué tan bien funcionó la randomización comparando la similitud de los grupos de tratamiento y de control.

El siguiente paso es recolectar una *encuesta de seguimiento* después de la intervención. De acuerdo al tipo de intervención, será importante considerar cuándo es más oportuno recolectar estos datos. Sobre todo, si es una intervención que necesita tiempo para mostrar impactos o si los impactos se desvanecen en el tiempo.

Finalmente, si quisiéramos evaluar cómo cambian los impactos a través del tiempo, es posible *continuar recolectando datos* sobre la misma muestra en el tiempo. Para esto, es muy importante que al recolectar la primera y segunda ronda de encuestas tengamos un identificador individual y toda la información necesaria para seguir a los hogares o individuos en el tiempo.

**Diapositiva 7**

En resumen, revisar la cadena de resultados ayuda a determinar qué tipo de información recolectar. Las consideraciones de costos y tipo de estudio determinan qué datos se recolectarán. Los indicadores recolectados deben ser SMART y es importante saber dónde y cuándo recolectar la información.

**Diapositiva 8**

Si quieres saber más sobre este tema te recomienda revisar estas referencias.

**Diapositiva 9**

Esto es todo. Espero te sea útil.