

PUBLICO

Diseño de El paso a paso I: Aspectos Operativos

Oficina de Planificación Estratégica y Efectividad en el Desarrollo Semana de Diseño de Evaluación de Impacto

De conformidad con la Política de Acceso a Información, esta presentación está sujeta a divulgación pública.

Copyright © 2018 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 nocomercial 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.

El contenido de esta presentación está basado parcialmente en Gertler et al (2011).



Agenda

- 1. Diseño Evaluación de Impacto
- 2. Recolección de datos
- 3. Potencia estadística y tamaño de muestra

Cronograma típico de evaluaciones de impacto

PROYECTO:



E.I. PROSPECTIVA:

DISEÑO	IMPLEMENTACIÓN	ANÁLISIS	DISEMIN ACIÓN	
			7.0.0	

Cronograma típico de evaluaciones de impacto

PROYECTO:

DISEÑO Y PLANIFICACION

IMPLEMENTACIÓN

CIERRE

E.I. PROSPECTIVA:

DISEÑO

IMPLEMENTACIÓN

ANÁLISIS

DISEMIN ACIÓN

DISEÑO:

- 1. Diagnóstico
- 2. Lógica vertical
- 3. Preguntas de evaluación
- 4. Metodología @
- 5. Cálculos de poder y muestra @

- 7. Fuentes de datos
- 8. Cronograma @
- 9. Presupuesto ወ
- 10. Identificar el equipo técnico @
- 11. Documentar en nota de concepto @
- 12. Presentar en IRB / Comité de Ética @

Diseño de una evaluación de impacto

Recurso muy útil:

PORTAL DE EVALUACIÓN DEL BID

www.iadb.org/portalevaluacion

EFECTIVIDAD PARA MEJORAR VIDAS

MIDIENDO RESULTADOS E IMPACTO



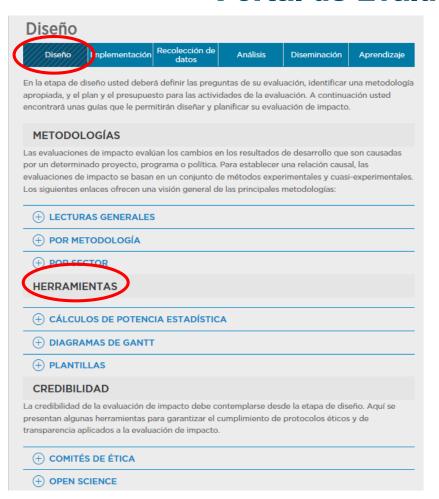


Bienvenido al portal de evaluación de impacto

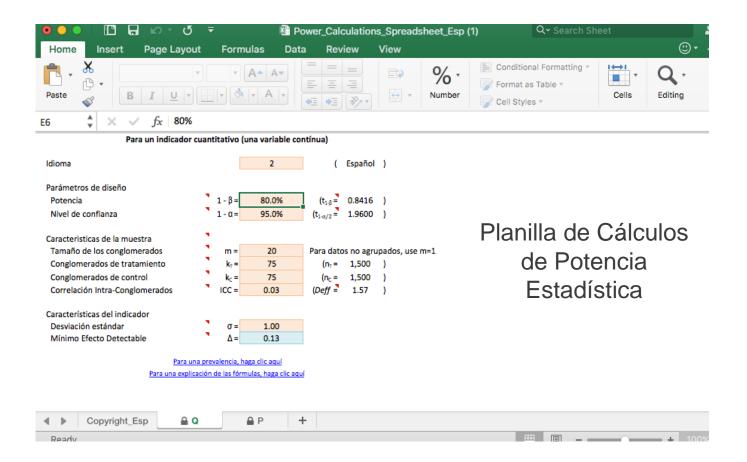
Este portal contiene herramientas para implementar evaluaciones de impacto. El contenido se organiza según el ciclo de actividades necesarias para completar una evaluación satisfactoria.

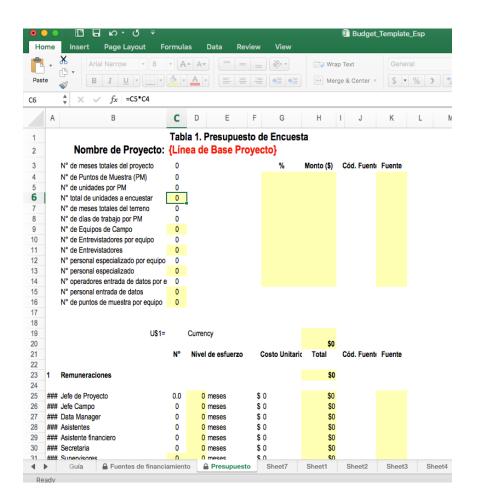


Diseño	Implementación	Recolección de datos	Análisis	Diseminación	Aprendizaje
Métodos Cálculo de potencia Lista de verificación Presupuesto Cronograma	Expresión de interés Términos de referencia Evaluadores	Cuestionarios Captura de datos Manuales del encuestador	Códigos Programas	Informes Resumen de políticas Panorama de efectividad en el desarrollo Inventario de evaluaciones	Próximos eventos Talleres previos Materiales decapacitación Lecciones virtuales

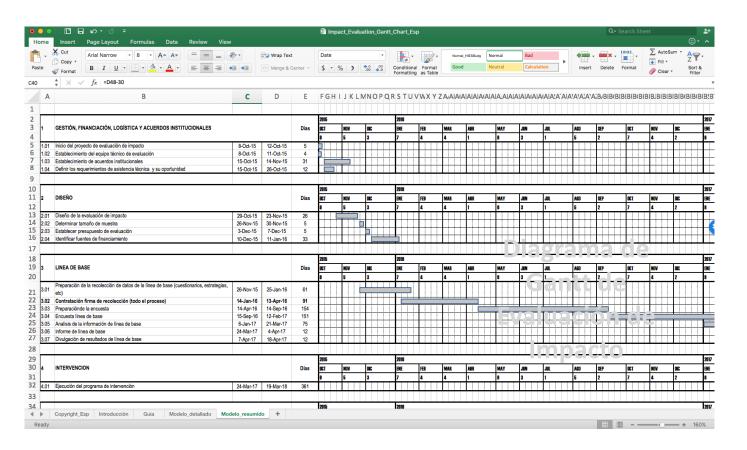


Cuenta con herramientas
 útiles para el diseño,
 implementación, análisis y
 diseminación de una
 Evalución de Impacto





Presupuesto Evaluación de Impacto



<u>Diseño</u>	Implementación	Recolección de datos	Análisis	Diseminación	Aprendizaje
 Métodos Cálculo de potencia Lista de verificación Presupuesto Cronograma 	 Expresión de interés Términos de referencia Evaluadores 	 Cuestionarios Captura de datos Manuales del encuestador 	Códigos Programas	 Informes Resumen de políticas Panorama de efectividad en el desarrollo Inventario de evaluaciones 	 Próximos eventos Talleres previos Materiales decapacitación Lecciones virtuales





Midiendo el Impacto: Pareamiento

Rosangela Bando

Oficina de Planificación Estratégica y Efectividad en el Desarrollo Washington, DC

El contenido de esta presentación representa la opinión del autor y no necesariamente del Banco Inter Americano de Desarrollo. Esta presentación incluye ilustraciones del libro "impact Evaluation in Practice" por Gertler, Martinez, Premand, Rawlings y Vermeersch (2010).



Diseño	Implementación	Recolección de datos	Análisis	Diseminación	Aprendizaje
Métodos Cálculo de potoncia Lista de verificación Presupuesto Cronograma	Expresión de interés Términos de referencia Evaluadores	Cuestionarios Captura de datos Manuales del encuestador	Códigos Programas	Informes Resumen de politicas Panorama de efectividad en el desarrollo Inventario de evaluaciones	Próximos eventos Talleres previos Materiales decapacitación Lecciones virtuales

"Check List" de Evaluación de Impacto

	Impact Evaluation Checklist from A to Z	Institution/ person in charge	Completed (Yes/No)	Resources
signing	an Impact Evaluation			
1.1	Diagnostic			
	1.1.1 Identify and quantify the development problem of interest			
	1.1.2 Identify context factors in the problem diagnostic			
	1.1.3 Describe the target population and sub-populations of interest			
	1.1.4 Review the literature related to the problem and target population			
	1.1.5 Review the literature for evidence based interventions (both for internal and external validity) used to address this problem			
1.2	Identification of a cost-effective intervention to address the problem			
	1.2.1 State the purpose of the intervention and rationale within national context			
	1.2.2 Identify intervention components/activities			
	1.2.3 Identify intervention location(s)			
	1.2.4 Conduct economic analysis (cost-benefit and/or cost-effectiveness) to determine economic viability of proposed intervention			Economic analysis webpape
	1.2.5 Define the intended use of results by policy and decision makers			
1.3	Logical Framework			
	1.3.1 Define the project components to be impact-evaluated			
	1.3.2 Spell out the evaluation questions and hypotheses			
	1.3.3 Identify SMART primary and secondary outcome indicators (program impacts and results)			
	1.3.4 Calculate summary statistics (means, prevalence, standard deviation) for primary and secondary outcome indicators			
	1.3.5 Map the program's logic (through a results matrix, theoretical model, or other format)			
	1.3.6 Identify data collection means (questionnaire, samples, etc.)			
1.4	Identify an evaluation methodology			Impact evaluation methods
	1.4.1 Define whether the evaluation will be experimental or quasi-experimental			
	1.4.2 Identify the optimal unit of treatment assignment (individual, household, village, school, health clinic, firm, farm, etc.)			
	1.4.3 Identify unit of analysis and of cluster (if applicable)			
	1.4.4 Identify whether treatment can be blinded or double blinded			
	1.4.5 Document randomization or treatment assignment procedure			
	1.4.6 Document treatment and comparison groups			
	1.4.7 Identify complementary evaluation strategies for the project including mixed methods, process evaluations, etc.			
1.5	Conduct power calculations (significance = 0.05 and power = 0.8 or 0.9)			Power calculations tool
	1.5.1 Identify minimum detectable effect size (MDE) using CBA, existing literature, simulation analysis, models, etc.			
	1.5.2 Calculate intra-cluster correlation using existing data (if applicable)			
	1.5.3 Calculate sample sizes (number of units and clusters) for each sub-population of interest that identify the proposed MDE			
	1.5.4 Adjust (increase) sample size to account for sample attrition			
1.6	Sample Selection			
	1.6.1 Identify sample framework including sampling universe			
	1.6.2 Identify stratification criteria			
	1.6.3 Determine number of stages for sample selection			
	1.6.4 Determine sample weighting properties (self-weighted or with sample weights)			

Agenda

- Diseño Evaluación de Impacto
- 2. Recolección de datos
- 3. Potencia estadística y tamaño de muestra

Definición fundamental: fuente de datos

- Cuáles son las fuentes de datos para la evaluación de impacto tiene efecto fundamental sobre
 - Diseño de la evaluación
 - Presupuesto
 - Tiempos de la evaluación
- Aspecto a veces más importante:
 ¿hay que recolectar datos?
 - Puede ser el componente más caro de la evaluación



Definición fundamental: fuente de datos

Fuentes posibles de datos para evaluaciones de impacto

- 1. Encuestas y censos (hogares, firmas) existentes
 - Suele ser la fuente más barata y fácil
 - A veces es difícil que se ajusten a las necesidades de información de la evaluación (temporal, geográfica, temas cubiertos, etc.)
- 2. Administrativas → registros de contribuyentes, beneficiarios, firmas, pacientes, etc.
 - No es utilizada suficientemente → requiere en general acuerdos y convenios con entidades públicas
 - Permite seguimiento continuo, de largo plazo y a un gran número de unidades (individuos, firmas, etc.)

Definición fundamental: fuente de datos

Fuentes posibles de datos para evaluaciones de impacto

- 3. Recolección primaria → implica realizar trabajo en campo (personas, hogares, firmas, comunidades)
 - Requiere mucho tiempo de planificación
 - Puede ser muy costosa
 - Permite obtener la información exactamente necesaria para la evaluación

Recolección primaria de datos implica definir muchos aspectos

- ¿Cuándo medir? ¿Deberíamos tomar Línea de Base?
- Seguimientos:
 - Se debe esperar lo suficiente para que los impactos se materialicen
 - Recolectar buenos datos de contacto en la línea base permite tomar varios seguimientos. ¿Hay recursos para eso?
- Si la selección de beneficiarios no se ha realizado, y es posible hacerla mediante una lotería → se puede aplicar el diseño experimental
- Hay que tener en cuenta los potenciales problemas de aleatorización

Pregunta clave: ¿Cómo balancear los costos y calidad de una medición en una evaluación de impacto?

- ¿Cuáles son los parámetros de calidad que se deben resguardar?
- ¿Cuáles son los elementos clave de un presupuesto de evaluación?
- ¿Cuál es la mezcla de calidad y costos que se busca al contratar una empresa encuestadora?

Algunos principios:

"Basura que entra, basura que sale"

- Cabalidad: Datos completos levantados de la totalidad de la población de estudio.
 - ✓ Violación: Encuestas con secciones no llenadas.
- Precisión: Datos bien especificados, capturan correctamente los parámetros de interés.
 - ✓ Violación: Preguntas conformadas con lenguaje confuso.
- Integridad: Datos obtenidos según los protocolos de aplicación y procedimientos operativos.
 - ✓ Violación: Aplicación de preguntas de encuesta en el orden incorrecto.

Algunos principios:

- Veracidad: Datos que expresan verazmente la información de la población de estudio.
 - ✓ Violación: Primado, falsificación de datos.
- **Homogeneidad:** Datos levantados por los mismos procesos y bajo las mismos protocolos de la población de estudio.
 - √ Violación: Encuestas tomadas con distintas instrucciones en distintas poblaciones.

¡Los errores de calidad son particularmente graves cuando son sistemáticos y/o no distribuidos de igual manera entre tratamiento y control!

- Algunas recomendaciones:
 - La planificación del terreno y diseño de cuestionario toma tiempo!
 No se confíe. Trabaje con anticipación.
 - Pilotear los instrumentos es fundamental.
 - ✓ Para corregir formulación de preguntas.
 - ✓ Para calibrar escalas de respuestas.
 - ✓ No haga 1 piloto, haga varios.
 - Buena coordinación con la muestra ayuda a reducir la "deserción".
 Esto es clave!
 - Monitorear el levantamiento de datos es clave.
 - Recolectar buenos datos es caro, pero "vale la pena".

Agenda

- Diseño Evaluación de Impacto
- 2. Recolección de datos
- 3. Potencia estadística y tamaño de muestra

Diseño de la evaluación y recolección de datos

- El diseño de la evaluación puede tener incidencia sobre la implementación del programa y sobre la recolección de datos
- Debe incluir un análisis de:
 - Marco Muestral
 - Estratificación
 - Cálculo de Potencia (o Poder)
- Para evaluaciones aleatorias, además
 - Unidad de Aleatorización

Marco Muestral

- El objetivo de tener un marco muestral es
 - Identificar personas o unidades que componen la población de estudio
 - Deben estar accesibles para aplicar el proceso de selección.
- Se crea una lista de elegibles con base en los criterios que se quieren evaluar en el programa.
- Marco muestral ideal : Toda la población, pero por razones de costo sólo se considera una parte.
- Condiciona la interpretación de los resultados

Marco Muestral: Ejemplo CLIP - Colombian Laundry Initiative Program

Programa

- Lavadoras (lavarropas) subsidiadas a familias de bajos recursos a través de lotería pública.
- Entrenamiento en uso de lavadora (uso apropiado, ahorro agua, promover participación masculina).

Impacto esperado

- ¿Tecnología de producción doméstica genera cambios en
- Uso del tiempo?
- Calidad de vida: laboral, ingresos, salud?
- Niños: nutrición, salud, desarrollo cognitivo y socioemocional?
- Roles de género, empoderamiento?

Marco Muestral: Ejemplo

Definición de los grupos de tratamiento y control CLIP – Colombian Laundry Initiative Program

- Hogares de bajos recursos:
 - Tres estratos socio-económicos más bajos
 - Zonas marginales de Bogotá
 - Familias biparentales con niños menores de 8 años

Experimentos: Unidad de Aleatorización

- Seleccionar según el tipo de programa:
 - Individual/Hogar
 - Vehículo
 - Calle/Camino
 - Pueblo/Comunidad
 - o Distrito/Municipio/Región



Tener en cuenta:

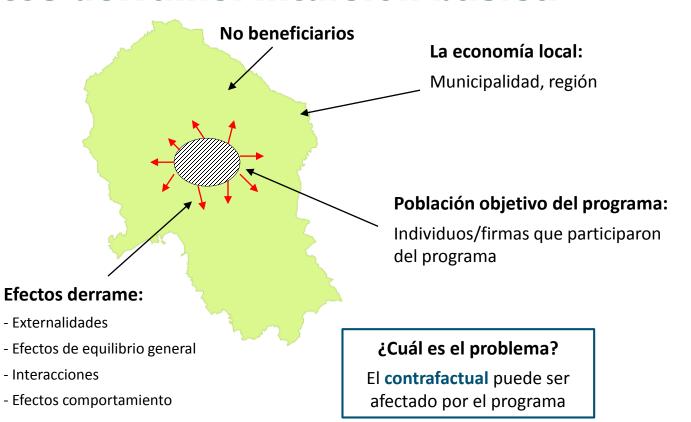
- Es necesario un número "suficientemente grande" de unidades para detectar el impacto mínimo deseado:
 Potencia estadística.
- Efecto Derrame (Ejemplo: campaña de desparacitación)
- Costos operativos y de encuestas

Efectos derrame (spillovers)

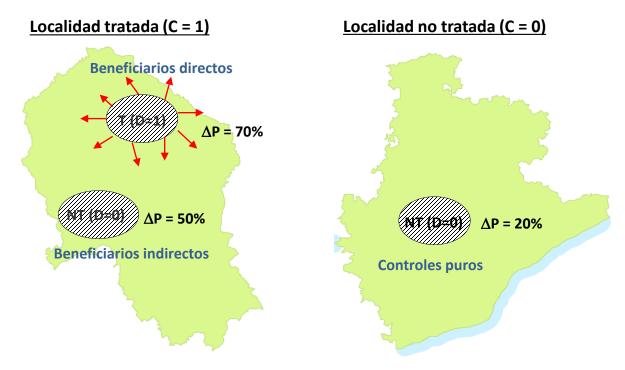
¿Porqué nos importan los efectos derrame?

- Si no se toman en cuenta pueden invalidar los resultados de una evaluación de impacto
- 2. Si son de magnitud importante, tenemos que medirlos

Efectos derrame: intuición básica



Efectos derrame: intuición básica



Efecto Tratamiento Directo = $\Delta P[C=1,D=1] - \Delta P[C=0,D=0] = (70\% - 20\%) = 50\%$ Efecto Tratamiento Indirecto = $\Delta P[C=1,D=0] - \Delta P[C=0,D=0] = (50\% - 20\%) = 30\%$

Efectos derrame: intuición básica



Efecto Tratamiento Directo = $\Delta P[C=1,D=1] - \Delta P[C=0,D=0] = (70\% - 20\%) = 50\%$ Efecto Tratamiento Indirecto = $\Delta P[C=1,D=0] - \Delta P[C=0,D=0] = (50\% - 20\%) = 30\%$

Experimentos: Unidad de Aleatorización

La selección de los grupos Ganadores y Perdedores (G&P) de la lotería - CLIP

- Se aleatoriza a nivel de hogar.
- Lotería pública, para lograr más transparencia en las comunidades.
- Pero, puede haber efectos derrame: hogares pueden prestar la lavadora a los que no ganaron.

Aleatorización

CLIP

- ¿Por qué variables estratificar aleatorización?
 - Los efectos externos se podrían dar a nivel de comunidad
 - Estratificar por comunidades
 - Distancia geográfica
 - Información cualitativa en barrios 'grandes' para determinar cuáles son los conglomerados (clusters)
 - Dado que es experimental: aleatorizar la intensidad del tratamiento para poder medir efectos derrame externos

Balance entre grupos

Es relevante revisar estadísticamente si los grupos de la evaluación están balanceados <u>antes</u> de la intervención.

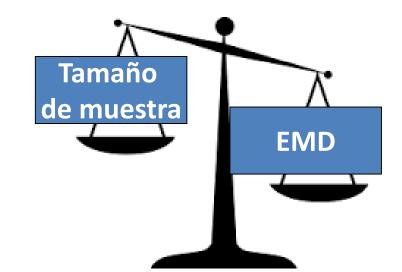
- ¿En qué dimensiones?
 - ✓ En todas las variables con que se cuente en la línea de base →
 qué puedan estar relacionadas con la intervención
- ¿Entre qué grupos?
 - ✓ Los grupos de la evaluación.
 - ✓ Puede ser relevante mostrar que los "desertores" son iguales entre grupos.

¿De que tamaño debe ser la muestra?

Depende principalmente de la magnitud de impacto que se quiera detectar:

EMD = Efecto Mínimo Detectable

- Queremos que el EMD sea pequeño...
- Pero esto requiere una muestra grande (mayores costos)



Ejemplo

Cuál es el efecto mínimo que se puede detectar (EMD) con esta muestra?

Caso: Impacto de apoyos productivos

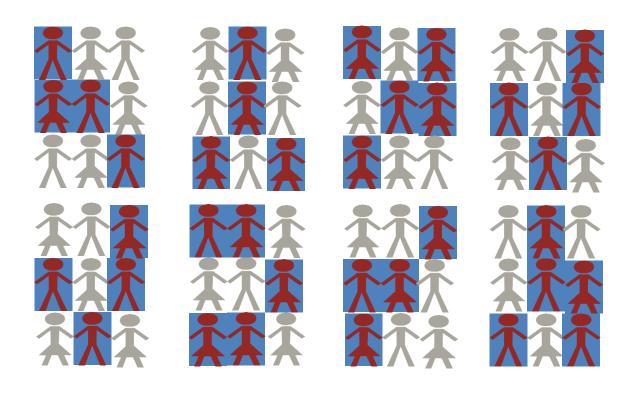
- 8 comunidades (conglomerados)
- 9 productores por comunidad
- Desviación estándar:
 - 1.000 pesos en ingreso anual

¿Aleatorizar a nivel de Individuo o Conglomerado?



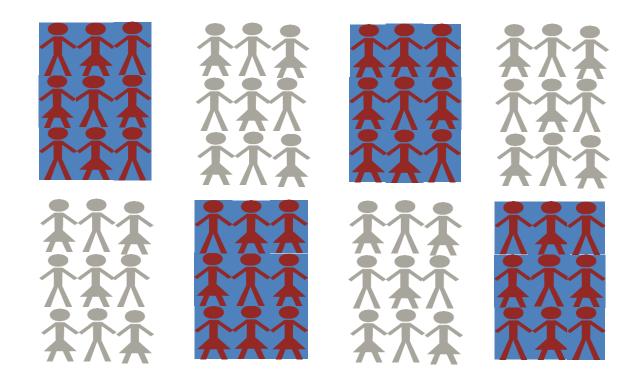
Aleatorización a nivel individual

Mayor poder pero efectos derrame potenciales



Aleatorización a nivel de conglomerado

Menor potencial de efectos derrame pero menos poder!



Por Conglomerado

Para un indicador cuantitativo (una variable contínua)

٦

Idioma

2 (Español)

Para datos no agrupados, use m=1

Parámetros de diseño

Potencia

Nivel de confianza

Características de la muestra
Tamaño de los conglomerados
Conglomerados de tratamiento
Conglomerados de control
Correlación Intra-Conglomerados

Características del indicador Desviación estándar Mínimo Efecto Detectable

$$1 - \beta =$$
 80.0% $(t_{1-\beta} = 0.8416)$
 $1 - \alpha =$ 95.0% $(t_{1-\alpha/2} = 1.9600)$

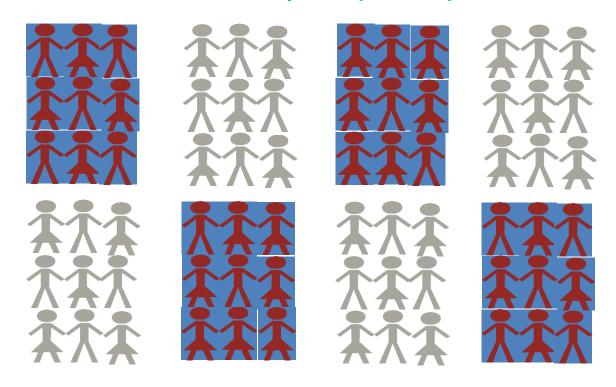
m = 9 k_T = 4

$$k_T = 4$$
 $(n_T = 36)$
 $k_C = 4$ $(n_C = 36)$
 $ICC = 0.2$ $(Deff = 2.6)$

 $\sigma = 1,000.00$ $\Delta = 1,064.77$

Por Conglomerado

EMD = **1060** pesos (CIC=0.2)



Por Individuo

Para un indicador cuantitativo (una variable contínua)

4

Idioma

2 (Español)

Parámetros de diseño

Potencia

Nivel de confianza

Caracteristicas de la muestra Tamaño de los conglomerados Conglomerados de tratamiento Conglomerados de control

Características del indicador Desviación estándar Mínimo Efecto Detectable

$$1 - \beta = 80.0\%$$
 $(t_{1-\beta} = 0.8416)$ $(t_{1-\alpha/2} = 1.9600)$

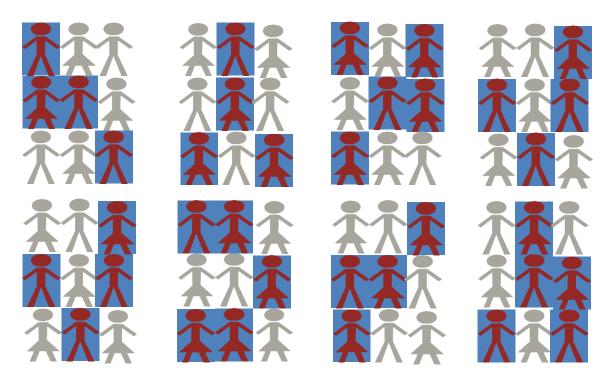
m = 1 $k_T = 36$ $k_C = 36$

Para datos no agrupados, use m=1 $(n_T = 36)$ $(n_C = 36)$

$$\sigma = 1,000.00$$
 $\Delta = 660.34$

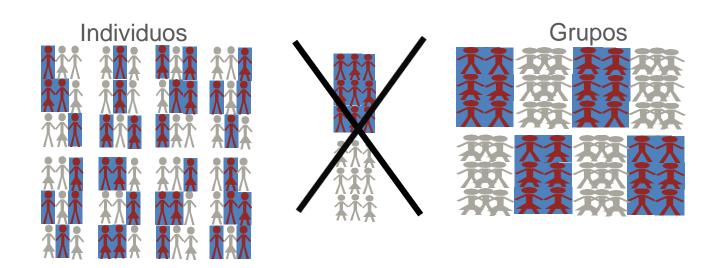
Por Individuo

EMD = 660 pesos



Entre más grande sea el grupo que se aleatoriza, más grande es el estudio

El número (de muestra) más importante es el número de unidades que se aleatorizan



Resumen

Diseño de Evaluación de Impacto...

- Implica bastante trabajo y seguir un protocolo
- Exige el uso de recursos → sobre todo si requiere relevamiento de datos!
 - Cálculos de potencia determinan costo de El
- Exiger ser creativo
- Y toma tiempo

Pero

 General evidencia relevante para mejorar la toma de decisiones y la gestion efectiva de los recursos!

