



Estudio Prospectivo Ciudad de México (MCPS)

Guía del usuario del *Data Showcase*: Primeros pasos

1 Introducción

El Estudio Prospectivo de la Ciudad de México (MCPS) se estableció en 1998 e incluye datos sobre más de 150,000 adultos de mediana edad. El estudio se está realizando en colaboración con la Universidad Nacional Autónoma de México en la Ciudad de México y ha recibido financiamiento de la Secretaría de Salud de México, el Consejo Nacional de Ciencia y Tecnología de México (Consejo Nacional de Ciencia y Tecnología), la *British Heart Foundation*, el Consejo de Investigación Médica del Reino Unido y el Wellcome Trust del Reino Unido.

El Showcase tiene como objetivo presentar los datos disponibles del estudio de una manera completa y concisa, y proporcionar información adicional relevante para los investigadores que estén considerando postularse para utilizar el recurso.

Esta guía del usuario está diseñada para brindar una breve descripción general de los datos de MCPS y proporciona algunas instrucciones sobre cómo navegar por el sistema.

Sugerencias e información para nuevos usuarios:

- Lea la información de antecedentes sobre MCPS y los detalles sobre los procedimientos de acceso a datos que se pueden encontrar en el sitio web (<https://www.ctsu.ox.ac.uk/research/prospective-blood-based-study-of-150-000-individuos-en-mexico>)
- Tenga a la mano una copia impresa de esta guía del usuario y tómese el tiempo para familiarizarse con la estructura de Showcase, la documentación adjunta y las descripciones provistas para cada campo de datos antes de completar una solicitud preliminar para acceder al recurso.

Si encuentra problemas o fallas, envíe un correo electrónico a mcps@ndph.ox.ac.uk

2 Datos incluidos en MCPS

2.1 Datos recopilados en la visita inicial (encuesta basal)

Entre 1998 y 2004, el *Clinical Trial Service Unit & Epidemiological Studies Unit* (CTSU) de la Universidad de Oxford, en colaboración con la Secretaría de Salud de México, establecieron un estudio en la Ciudad de México, en el que participaron más de 150 000 adultos de mediana edad (incluidos 100 000 mujeres y 50.000 hombres) proporcionando información sobre su estilo de vida y su historial de enfermedades. También se registraron medidas antropométricas (incluido el peso, la circunferencia de la cintura y la cadera, la presión arterial) y se les tomó una muestra de sangre.

2.2 Datos recopilados en la visita de reencuesta

Una re- encuesta de 10 000 participantes sobrevivientes (2015 - 2019) capturó cómo los estilos de vida, las mediciones físicas y biológicas y los tratamientos para enfermedades (por ejemplo, diabetes) han cambiado con el tiempo. La re-encuesta también incluyó varias "mejoras" (como la bioimpedancia y la recolección de una muestra de orina).

2.3 Datos de seguimiento vinculados

Ahora se realiza un seguimiento de la mortalidad de todos los participantes a través de la vinculación con las bases de datos nacionales de mortalidad de México; para enero de 2018, se confirmó la muerte de más de 20.000 participantes del estudio.

2.4 Solicitud de datos y futura disponibilidad de datos

La política de acceso a datos y muestras (disponible en el sitio web, tanto en [inglés](#) como en [español](#)) brinda detalles completos sobre cómo solicitar acceso a los datos del estudio de MCPS, así como información sobre actualizaciones y el cronograma de publicación de datos.

Puede enviarnos un correo electrónico a mcps-access@ndph.ox.ac.uk con cualquier consulta específica sobre el acceso a los datos o sobre el estudio en general.

3 Búsqueda de datos en Showcase

BROWSE: use esta opción para navegar a través de categorías jerárquicas y subcategorías de interés a campos de datos (es decir, variables) de interés. **Esta será la herramienta más apropiada para la mayoría de los investigadores que deseen encontrar y seleccionar datos para su aplicación para utilizar el Recurso.**

CATEGORÍAS: Los campos de datos relacionados se han agrupado en categorías, como por ejemplo *Datos de muestras de sangre* o *Consumo de tabaco*. Puede encontrarlos en el listado de **CATEGORÍAS** en la **sección CATÁLOGOS** (<https://datashare.ndph.ox.ac.uk/mexico/cats.cgi>).

Puede encontrar el resto de los datos que necesita a través del buscador:

SEARCH : La búsqueda predeterminada es una búsqueda de texto del nombre del campo de datos, sus notas y codificaciones de datos. Ingresar un valor numérico en el cuadro de búsqueda devolverá el campo de datos con esa ID. La función **de búsqueda de campos de datos** también le permite realizar una búsqueda utilizando criterios específicos basados en el tipo de campo de datos (consulte la Sección 5 para obtener más detalles).

Las búsquedas de texto de codificación de datos, categorías, recursos y conjuntos de datos se pueden realizar seleccionando el botón de tipo de búsqueda correspondiente.

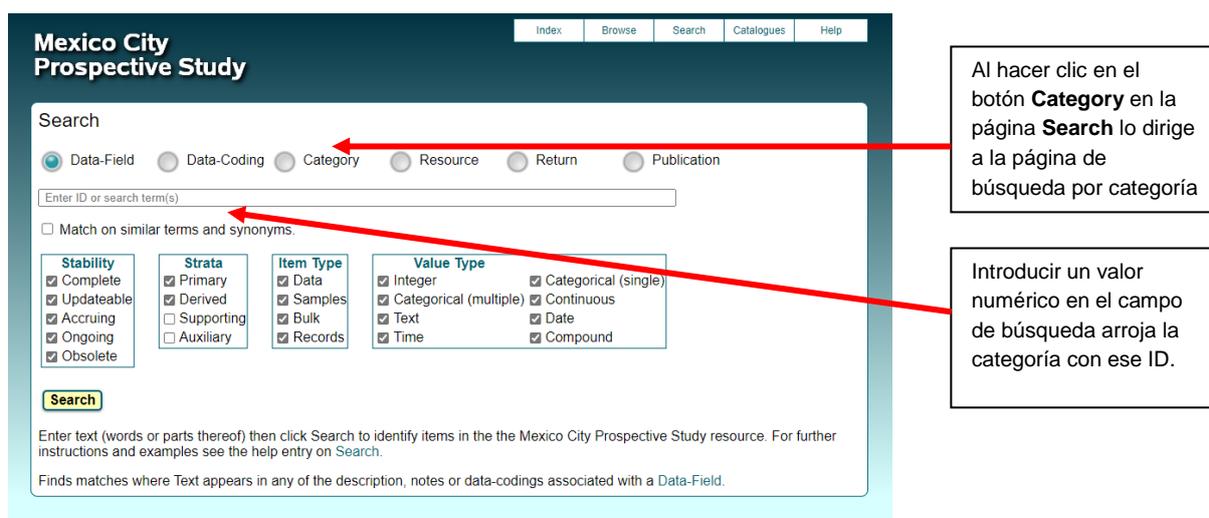


Figura 1. Ilustración de la búsqueda por categorías (Search)

Consulte la pestaña **HELP** en la página de búsqueda para obtener más detalles sobre cómo realizar búsquedas de texto.

Puede encontrar una lista completa de campos de datos, categorías y documentos en la pestaña **CATALOGUES**.

4 Categorías y subcategorías de datos

Los datos están organizados en una estructura de árbol, accesible a través de **BROWSE**, con las principales categorías basadas en el origen de la recopilación de datos (Figura 2). Éstos incluyen:

- Evaluación de referencia
- Seguimiento de la mortalidad
- Evaluación de reencuesta

Consulte la página de **HELP** en 'Examinar' para obtener más detalles.

Los campos de datos y/o las subcategorías dentro de una categoría en particular se revelan al hacer clic en **Category ID** o en **Description** en la tabla que se muestra (Figura 3).

Mexico City Prospective Study

Index Browse Search Catalogues Help

Browse by Primary Category of Origin

Category	Items
Baseline: Assessment	90
Resurvey: Assessment	127
Mortality follow-up	8

Top Level
Level 1
Level 2
Level 3

Summary generated 3 May 2022

La columna **Items** enlista el número de campos de datos en cada categoría (y en sus sub-categorías).

Hacer clic en los botones **Level** es una forma fácil de brincar hacia un nivel más detallado del árbol

El botón **Help** provee más información sobre los **items**, según la lista en el **Glossary** (al final de la página **Help**).

Figura 2. Ilustración de la estructura de árbol a través de **BROWSE**

Mexico City Prospective Study

Index Browse Search Catalogues Help

Category 3
Baseline: Lifestyle characteristics - Baseline: Assessment

Description
This category contains information, collected from the baseline questionnaire, on lifestyle characteristics, subdivided into categories.

Notes 3 Sub-Categories 4 Data-Fields 1 Parent Category

Category ID	Description	Items
4	Baseline: Smoking	12
5	Baseline: Alcohol consumption	3
6	Baseline: Physical activity	3

Hacer clic en **Category ID** o en **Description** lo dirige a las subcategorías y/o los campos de datos contenidos dentro de dicha categoría.

Figura 3. Ilustración de subcategorías dentro de la categoría 'Características del estilo de vida'

La estructura de árbol asigna campos de datos a una sola ubicación y actualmente no tiene referencias cruzadas. Por lo tanto, es importante buscar en todas las partes del árbol que puedan contener campos de datos relevantes para su(s) pregunta(s) de investigación. En general, no debe confiar en la función de **SEARCH** para encontrar todos los campos de relevancia para un tema en particular.

5 Información del campo de datos

El panel en la mitad superior de la pantalla del campo de datos proporciona una descripción breve y la ubicación de la categoría del campo de datos dentro de la estructura de árbol (Figura 4). También incluye información técnica más detallada sobre cada campo de datos. Esto incluye información sobre:

- **Participantes:** el número de participantes que tienen el elemento de datos
- **Recuento de elementos:** el número de elementos de datos disponibles
- **Estabilidad:** si el campo de datos está completo o cambia con el tiempo
- **Tipo de valor:** el formato y las unidades del campo de datos
- **Tipo de elemento:** si el campo de datos es un punto de datos simple, si se relaciona con un inventario de muestras biológicas o si es un objeto de datos grande
- **Estratos:** la probable relevancia para los investigadores del campo de datos
- **Sexado:** si el campo de datos está disponible para ambos sexos
- **Instancias:** cuántas veces los participantes han realizado esta medición
- **Array:** si hay varios elementos de datos para cada instancia. Por ejemplo, la Figura 4 muestra que los datos sobre la presión arterial diastólica se presentan en una matriz con 3 valores por medida (porque la medición se realizó tres veces). Consulte la página de **HELP** para obtener más detalles.

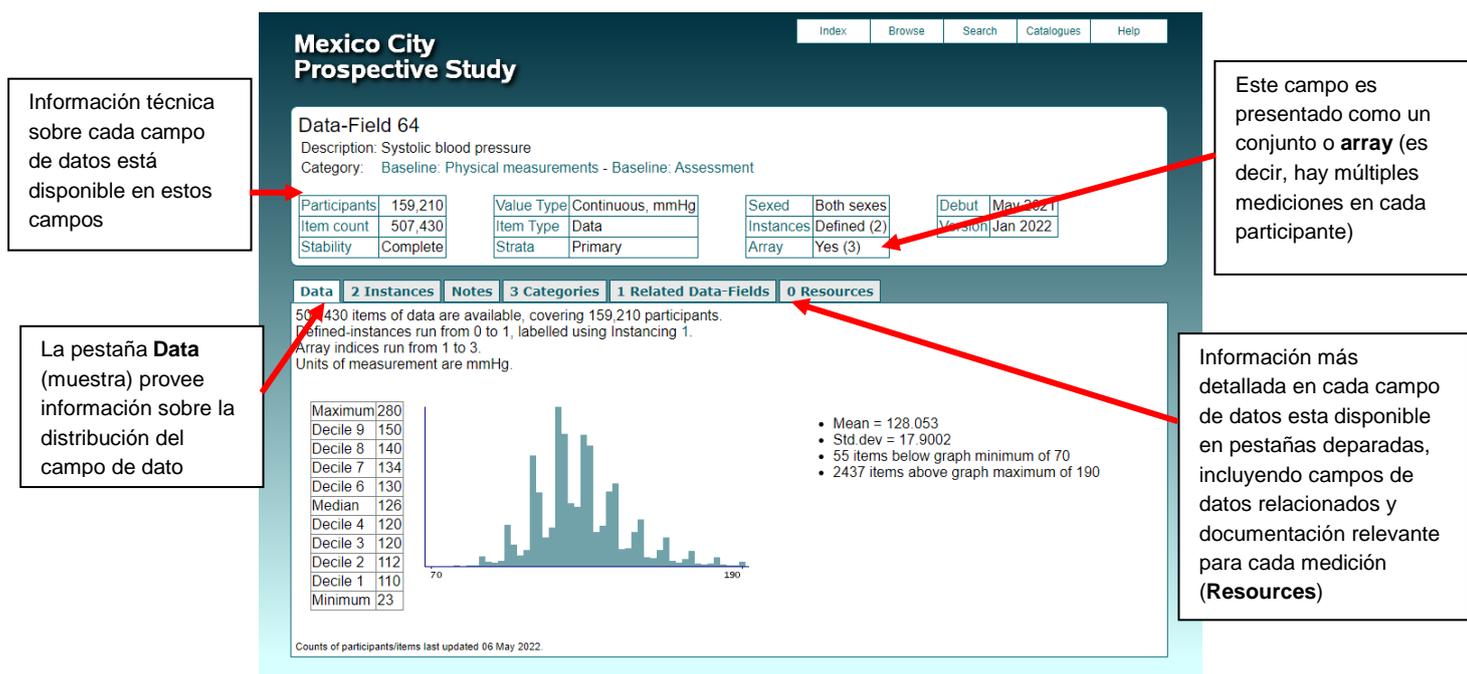


Figura 4. Ilustración de una página de campo de datos

La distribución univariada de cada campo de datos se presenta en formato gráfico o tabular (o ambos) en la pestaña **Data** (Figura 4). Los datos no se presentan si son de texto libre u otros elementos de datos inadecuados.

La pestaña **Instances** proporciona la distribución univariada de cada campo de datos en cada instancia (p. ej., para el campo de datos 64, los datos se presentan por separado para la evaluación inicial de referencia (Instance 0) y la evaluación de re-encuesta (Instance 1)).

La pestaña **Notes** incluye la descripción completa del campo de datos, junto con otros detalles.

La pestaña **Categories** enumera las categorías y subcategorías de las que el campo de datos es miembro. Esto también se muestra horizontalmente en el árbol de categorías, en la parte superior de la página.

La pestaña **Related Fields** enumera otros campos de datos con los que está relacionado el campo de datos actual. Por ejemplo, el campo de datos para “presión arterial diastólica” (ID: 65) está relacionado con el campo de datos 'presión arterial sistólica' (ID: 64).

La pestaña **Resources** contiene documentación explicativa relacionada con cada campo de datos, si corresponde.

6 Limpieza de datos

Los datos al inicio (encuesta basal 1998-2004) se ingresaron directamente en dispositivos portátiles y en la re-encuesta (2015-2019) en tabletas electrónicas (Samsung Galaxy 10.1). Todos los dispositivos se programaron para guiar a los trabajadores de campo (previamente capacitados) a través de los cuestionarios de la misma manera prescrita, omitiendo automáticamente las preguntas que no eran relevantes (por ejemplo, preguntas sobre antecedentes obstétricos para participantes masculinos). Los dispositivos utilizados para recolectar datos alertaron a los encuestadores sobre mediciones físicas moderadamente extremas e impidieron la entrada de valores altamente inverosímiles.

Al final de cada jornada laboral, los datos recogidos se cargaban automáticamente en la base de datos electrónica del estudio. Los procedimientos electrónicos redujeron al mínimo los costos de recopilación, procesamiento y verificación de datos, al tiempo que mejoraron la precisión y la integridad de los datos. Se utilizó un programa de verificación de la base de datos para identificar posibles errores de datos en la base de datos del estudio, los cuales fueron revisados regularmente por un equipo de monitoreo de datos y, cuando fue necesario, verificados por trabajadores de campo (p. ej., al volver a visitar a los participantes).

Los datos de Showcase se proporcionan 'tal cual' desde la base de datos electrónica del estudio; es decir, no ha habido validación adicional, limpieza de datos o cruce entre variables, además de lo mencionado anteriormente.

Los datos obtenidos a través de enlaces (es decir, registros de defunción) están sujetos a controles de validación y limpieza. Esto implica identificar ambigüedades en los datos, como códigos de clasificación clínica no válidos o discrepancias en los registros de los participantes.