

Consejo Asesor Técnico del Sistema Estadístico Nacional (CASEN) 2023-2025

Acta octava reunión ordinaria – Ayuda de memoria

Ciudad: Bogotá D.C

Lugar: Sesión virtual

Tema: Diseño y ejecución del Registro
Estadístico Base de Población - REBP

Hora: 09:00 a. m. a 11:00 a. m.

Fecha: 27/06/2025

Dependencia responsable: DCD-DIRPEN

Participantes

Miembros de la Sala Especializada Salud, Bienestar Social y Demografía

Lina María González Ballesteros, experta
asesora.

Fernando Urrea Giraldo, experto asesor.

Enrique Acosta, experto asesor.

Departamento Administrativo Nacional de Estadística (DANE)

Julieth Alejandra Solano Villa, directora
técnica de la Dirección de Regulación,
Planeación, Estandarización y Normalización
(DIRPEN) y secretaria técnica CASEN.

Javier Plazas Echeverri, director técnico de la
Dirección de Censos y Demografía (DCD).

Alexander Ernesto Paez Almonacid,
coordinador GIT Censos, Conteos y Operaciones
Estadísticas Especiales de Población

Jose Andrés Almanza Segura, DCD - GIT
Censos, Conteos y Operaciones Estadísticas
Especiales de Población

Steven Cifuentes Rugeles, DCD - GIT Censos,
Conteos y Operaciones Estadísticas Especiales
de Población

Cristian Duney Bermudez Quintero, DCD - GIT
Investigación y Desarrollo

Mariana Francisca Ospina Bohorquez,
coordinadora - GIT Proyecciones y análisis
demográfico

Alberto Antonio Moreno Córdoba, DCD - GIT
Proyecciones y análisis demográfico

Paula Andrea Amado Amado, DCD

German Leónidas Orjuela Borda, coordinador
GIT Planificación y Articulación Estadística.

Felipe Andres Martinez Fonseca, DIRPEN - GIT
Planificación y Articulación Estadística.

Mauricio Giovanni Valencia Amaya, DIRPEN -
GIT Planificación y Articulación Estadística.

Juan Sebastian Vásquez Acevedo, DIRPEN -
GIT Planificación y Articulación Estadística.

Carlos Eduardo Castro Zarate, DIRPEN - GIT
Planificación y Articulación Estadística.



Orden del día

Tiempo	Actividad	Responsable
09:00 a. m. a 09:05 a. m.	Instalación, verificación de quorum, permiso para grabar y registro fotográfico	Carlos Eduardo Castro Zarate, profesional especializado GIT PAE
09:05 a. m. a 09:10 a. m.	Apertura de la reunión	Julieth Solano, directora técnica DIRPEN y secretaria técnica CASEN
09:10 a. m. a 9:20 a. m.	Síntesis reunión anterior	Carlos Eduardo Castro Zarate, profesional especializado GIT PAE
09:20 a. m. a 10:00 a. m.	Línea: Fortalecimiento en la oferta de Operaciones Estadísticas y optimización de Registros Administrativos Título: Socialización técnica del Registro Base de Población - REBP fases de construcción, experiencias internacionales y herramientas tecnológicas.	Alexander Ernesto Paez Almonacid, coordinador GIT Censos, Conteos y Operaciones Estadísticas Especiales de Población Jose Andrés Almanza Segura, profesional especializado GIT Censos, Conteos y Operaciones Estadísticas Especiales de Población
10:00 a. m. a 10:40 a. m.	Realimentación por parte de los miembros de la sala	Miembros del CASEN. Modera: Carlos Eduardo Castro Zarate, profesional especializado GIT PAE
10:40 a. m. a 10:50 a. m.	Compromisos de la próxima reunión.	Julieth Solano, directora técnica DIRPEN y secretaria técnica CASEN.
10:50 a. m. a 11:00 a. m.	Conclusiones y cierre.	Julieth Solano, directora técnica DIRPEN y secretaria técnica CASEN.

Desarrollo

Objetivo

Socializar con los miembros expertos del CASEN las distintas fases de construcción del Registro Base de Población, incluyendo experiencias internacionales, junto con las herramientas tecnológicas que aseguran la robustez y seguridad de esta importante herramienta

1. Apertura

La sesión inició con la bienvenida por parte de Carlos Castro delegado de DIRPEN, quien explicó la dinámica de trabajo y la importancia de la instancia de coordinación del SEN. Se presentó al nuevo experto Enrique Acosta, especialista en demografía, quien se integró al equipo para aportar en el análisis de las temáticas trabajadas en la mesa estadística y de manera general se le indicó las dinámicas de trabajo de estas reuniones.

2. Síntesis de la reunión anterior

Carlos Castro de DIRPEN mencionó que en la sesión previa se consultaron a los expertos sobre los métodos e insumos para la producción y actualización de proyecciones demográficas del DANE. Se recibieron comentarios y recomendaciones que serán socializados al equipo de la dirección de censos y demografía para el fortalecimiento de las estadísticas producidas por el DANE. Se destacó la agenda proyectada para las siguientes sesiones que resume las fechas sugeridas para las actividades concertadas por las direcciones técnicas a fin de contar con la respectiva participación de los expertos asesores de esta sala de acuerdo con el siguiente cuadro resumen:

Tema	Objetivo	Fecha
Proyecciones de población	Validar con los expertos CASEN las metodologías de proyecciones a partir del aprovechamiento de RRAA	30 de mayo
Censos Mixtos	Actualización del Registro Estadístico Base de Población	27 de Junio
	Validar con los expertos CASEN las metodologías de censos mixtos	29 de agosto
	Análisis exploratorio del estado de las variables de discapacidad en los RRAA	26 de septiembre
Intersala	Sistema de Información Wayuu: Validar con los expertos CASEN los avances y resultados generados por el sistema (Indicadores, medición y difusión)	25 de Julio
Nota Estadística	Documentar información estadística basada en evidencias que ayuden a hacer análisis cuantitativos y cualitativos sobre	Agosto

diferentes fenómenos sociodemográficos (Salud Mental)

3. Presentación: Diseño y ejecución del Registro Estadístico Base de Población - REBP

La presentación del Registro Estadístico Base de Población (REPB) estuvo a cargo del equipo de la Dirección de Censos y Demografía del equipo del GIT Censos, Conteos y Operaciones Estadísticas Especiales de Población y abordó diversos aspectos clave relacionados con el diseño, desarrollo, normatividad, metodología, fuentes, calidad de datos y uso del REPB. A continuación, se detallan los contenidos tratados y una síntesis de cada punto:

A. Introducción y Contexto General

Se expusieron desafíos y retos asociados a los censos tradicionales, principalmente los altos costos y dificultades logísticas en Colombia por su geografía diversa. Se planteó la alternativa del uso de registros administrativos para disminuir costos, tiempos y desgaste de la fuente. Se destacó que el REPB no reemplaza el censo, pero es un apoyo importante para optimizar la producción estadística continua.

B. Normatividad Vigente

Se mencionaron leyes relevantes como la Ley 79 de 1993 art 5 que protege con reserva legal los datos entregados al DANE y no pueden ser divulgados al público o entidades para efectos legales o tributarios. Y el marco general de la Ley 2335 de 2023 que establece el marco jurídico general para la planificación, producción, difusión y administración de las estadísticas oficiales, con un componente muy importante en el fortalecimiento y aprovechamiento de registros administrativos, interoperabilidad, uso de fuentes alternativas y cultura estadística para la producción de información

C. Definición y Objetivos del REPB

El REPB es una base de datos integrada de registros administrativos con el objetivo de mejorar la calidad y coherencia de la información poblacional. Se orienta a producir datos estructurados y actualizados continuamente para apoyar la formulación de políticas públicas. Se aclaró que no es en sí misma una operación estadística, sino una herramienta para alimentar operaciones estadísticas oficiales.

D. Historia y Avances

El desarrollo del REPB inició hace más de 10 años con pruebas piloto en San Andrés (2015) y Jamundí (2016). En 2017 contó con apoyo del BID para lineamientos, consolidándose el registro en 2018. Entre 2019 y 2021 se actualizaron metodologías y se gestiona información para años posteriores (2022-2023).

E. Metodología y Documentación

Usando del modelo estadístico GSGBPM 5.1 para documentar fases del proceso estadístico. Se construyó una ficha técnica como documento principal y se trabaja en otros documentos de diseño, control de calidad, metadatos y accesibilidad. Se enfatizó el control y fortalecimiento de la calidad de datos a partir de registros primarios, con seguimiento y validación constante.

F. Fuentes de Información

Diversas fuentes primarias alimentan el registro: Registraduría Nacional (Registro civil de nacimiento, defunción, identificación), salud (BDUA, vacunación), educación (SIMAT, SNIES), seguridad social (PILA), migración (REMI), registro tributario, discapacidad, víctimas.

Se describió el proceso de solicitud de información, estandarización, homologación y clasificación para asegurar interoperabilidad.

G. Procesos Técnicos y Herramientas

Se utilizan herramientas automatizadas (Pentaho, SQL Server, Python) para la conexión, transformación, validación, y detección de duplicados. Se aplican técnicas de anonimización, supresión de variables sensibles y encriptación (hash) para protección de datos. Existe un repositorio institucional para sistematizar y documentar los scripts y procesos.

Se asigna un identificador estadístico único a cada persona basado en atributos claves (documento, nombre, fecha de nacimiento) mediante algoritmos determinísticos. Se mantiene estable el identificador para seguimiento longitudinal. La trazabilidad de la información se realiza mediante matrices de prioridad de fuentes, no a nivel granular en la base de datos.

H. Calidad y Control de Datos

Se aplican criterios jerárquicos para resolver inconsistencias entre fuentes según calidad, cobertura y actualización. Se reconoció la necesidad de mejorar la calidad en variables sensibles, como la información étnico-racial, con esfuerzos normativos

recientes para obligar su inclusión. Se resaltó la importancia de auditorías, control y sensibilización a los proveedores de datos, especialmente en el sector salud.

I. Limitaciones y Retos

Hay omisiones significativas especialmente en datos étnico-raciales, con pérdidas notablemente en población afrodescendiente. El REPB carece actualmente de información sociodemográfica completa y no captura estructuras de hogar ni vivienda como el censo. Se reconoció que el REPB actualmente no captura dinámicas demográficas como mortalidad o fecundidad, solo protocolos estáticos.

J. Usos y Accesibilidad

La difusión se realiza en formatos anonimizados con niveles de agregación que minimizan riesgos de re-identificación. Se cuenta con protocolos para acceso controlado a versiones ampliadas para investigaciones. El registro facilita interoperabilidad con otras plataformas y se trabaja en mejorar su usabilidad.

K. Perspectivas Futuras

Se está avanzando en la actualización y fortalecimiento normativo para que todos los registros administrativos incluyan información esencial. Se analiza la posibilidad de integrar información censal con el REPB para mejorar la calidad y cobertura. Se planifica un manual de auditoría de calidad y estrategias para involucrar a los tomadores de datos para aumentar la calidad. Se impulsa la construcción de censos mixtos para complementar el sistema y abordar omisiones.

La presentación evidenció que el REPB es un esfuerzo consolidado en Colombia para aprovechar registros administrativos y suplir las limitaciones de los censos tradicionales, buscando una información poblacional continua y actualizada. Se reconocen importantes avances metodológicos y técnicos para integrar fuentes diversas, asegurar calidad y proteger datos. No obstante, persisten retos en la cobertura, calidad (en particular en información étnico-racial), trazabilidad detallada y dinamismo de datos demográficos. El equipo se está enfocando en actualizar normativas, mejorar metodologías, fomentar la interoperabilidad y estudiar la integración con censos propios y de campo para robustecer este sistema. La discusión con expertos permitió identificar áreas prioritarias para profundizar y elevar la utilidad y confiabilidad del REPB a futuro

4. Realimentación miembros de la sala

Durante la sesión se plantearon varias preguntas importantes relacionadas con el Registro Estadístico Base de Población (REBP), las cuales fueron respondidas por el equipo de GIT Censos, Conteos y Operaciones Estadísticas Especiales de Población.

En la dinámica de preguntas e inquietudes se destacan las siguientes preguntas abordadas por los expertos y en general en la sesión:

Experta Lina María

Preguntas:

1. Preguntó sobre si las fuentes de datos son mayoritariamente bases secundarias o primarias.
 - Se aclaró que la fuente es primaria para la obtención de los registros, pero para el uso previo del registro es una fuente secundaria.
2. Preguntó sobre la ausencia de datos sociodemográficos en los registros administrativos comparados con el censo y si se tendría un gap permanente en este aspecto, especialmente en temáticas raciales y étnicas.
3. Preguntó si los registros actuales intentan homologar la información que captura el censo, pero que existe incertidumbre sobre si la suma de los registros es exacta o no.
4. Consultó si se ha pensado en diseñar un manual de auditoría de calidad de los datos.
5. Señaló que los tomadores de datos en salud (profesionales de salud) no están formados ni motivados para tomar datos de calidad, por lo que sugirió involucrar directamente a estas personas para mejorar la recolección.

Respuestas:

- Se explicó la diferencia entre fuentes primarias y secundarias.
- Se reconoció que hay gran omisión en temas como la información étnico-racial, pero que a partir del 2023 existe normativa para que la información se recoja mejor.
- Se comentó que los registros tratan de capturar la información del censo, pero se encontró inconsistencia entre las fuentes y que el censo no se usa para alimentar directamente el registro base, aunque podría ser útil para complementar y validar.
- Se coincidió en la importancia de diseñar un manual o estrategias para auditar y mejorar la calidad de los datos, especialmente con el sector salud.
- Se reconoció la importancia de sensibilizar a quienes toman los datos y que existen avances en algunas regiones con el sector salud.

Fernando Urrea

Preguntas / Comentarios:

1. Expresó preocupación por la cultura de registro en poblaciones indígenas y afro, donde muchos nacimientos y defunciones quedan en la memoria oral y no se registran formalmente.
2. Preguntó cómo se integra y utiliza la información de censos propios indígenas, y qué papel juegan en el registro estadístico base de población.

3. Aportó que el problema de omisión es más grave para población afro y menos para población indígena, donde el censo fue generalmente satisfactorio.
4. Propuso ampliar la interoperabilidad del registro con fuentes no oficiales para incluir censos propios y otras fuentes alternativas especialmente para pueblos indígenas.
5. Sugirió usar el censo como una medida transitoria para validar y contrastar la información del registro estadístico base, aunque a mediano y largo plazo el registro debería reemplazar al censo.

Respuestas / Comentarios:

- Se dijo que se están probando metodologías para identificar omisión con fuentes administrativas y censos propios.
- Se aclaró que el censo propio indígenas es una fuente que debe ser utilizada y mejorada mediante acuerdos con autoridades tradicionales.
- Se reconoció la importancia del censo para población indígena pero también su limitación para población afro.
- Se apoyó la idea de ampliar la interoperabilidad para incluir fuentes oficiales y no oficiales.
- Se concordó en que el registro base debe funcionar como sistema para el futuro reemplazo del censo, pero actualmente el censo es útil para validar la información.

Enrique Acosta

Preguntas / Comentarios:

1. Preguntó sobre la trazabilidad de los datos en el registro base, específicamente cómo se puede saber cuál fuente contribuye con cada dato y la necesidad de un protocolo para priorizar fuentes para variables como la información étnico-racial.
2. Señaló que hay un alto porcentaje de pérdida de información étnico-racial en el registro en comparación con el censo, y preguntó si el censo podrá alimentar o complementar el registro base.
3. Mencionó la importancia de evaluar las dinámicas demográficas (mortalidad, fecundidad) usando el registro base, tema que no está cubierto ahora pero debería implementarse en el futuro para tener un registro dinámico y funcional.
4. Propuso que el censo también podría servir para evaluar la calidad de las diferentes fuentes del registro.

Respuestas:

- El diseño actual del REBP no permite identificar la fuente de cada dato. Se puede determinar la jerarquía de las fuentes para cada variable a partir de la matriz, pero no el origen de cada dato.
- Se reconoció que actualmente no se usa el censo para alimentar el registro, pero es algo que podría reevaluarse para futuras versiones del registro.
- Se acepta la necesidad de mejorar la trazabilidad y la incorporación de variables adicionales como la información étnico-racial bajo normativa reciente.
- Se comentó que el registro base actualmente no puede hablar de dinámicas demográficas pero es un reto para desarrollos futuros.

- Se conjunto con la idea de usar el censo para evaluar y mejorar la calidad.

A continuación, se detallan las preguntas y respuestas abordadas:

1. **¿Cómo se resuelven las inconsistencias entre diferentes fuentes administrativas para una misma persona?**

Se aplican reglas jerárquicas basadas en la calidad, cobertura y actualización de cada fuente. Se cuenta con una matriz de prioridades donde se identifican las fuentes primarias y secundarias para cada tipo de información. Por ejemplo, la registraduría es la fuente primaria para datos básicos, mientras que para validar se usan fuentes secundarias como el BDUA.

2. **¿Se asigna un identificador estadístico único de persona y es estable en el tiempo?**

Sí, se utiliza un algoritmo determinístico que crea un identificador único de persona basado en atributos como número de documento, fragmentos de nombres, apellidos y fecha de nacimiento. Este identificador se pseudoanonimiza para agregar seguridad y permanece estable entre versiones del registro para permitir análisis longitudinales.

3. **¿Qué técnicas de anonimización se utilizan para evitar reidentificación, especialmente con variables sensibles?**

Principalmente se usan técnicas de supresión de identificadores críticos como nombres, apellidos y documentos, además de aplicar técnicas de encriptación (hash) para proteger la información sensible. Por ahora, no se ha realizado un análisis formal de riesgo de reidentificación, pero se contempla incorporar esta evaluación en futuras versiones.

4. **¿Qué herramientas y procesos están automatizados en el manejo del REBP?**

Se utilizan herramientas como Pentaho para la orquestación (conexión y transformación de datos) y SQL Server para procesamiento específico. También se emplea Python para tareas particulares. Los procesos automatizados incluyen detección de duplicados, identificación de registros comunes y depuración de datos, con algoritmos avanzados como Harold Winkler y Levenstein.

5. **¿Cómo se asegura la interoperabilidad semántica entre variables similares de diferentes fuentes?**

Se hace un proceso de homologación y estandarización de variables (sexo, tipo de documento, niveles educativos, etc.) para uniformar la codificación y facilitar los cruces. Además, se usa un modelo común de datos y un diccionario de equivalencias que se aplica a todas las fuentes administrativas.

6. **¿Cómo se realiza la difusión y acceso a la información del REBP?**

La difusión es a través de medios digitales con acceso controlado. Solo se publican formatos anonimizados y con variables agregadas para minimizar el riesgo de reidentificación,

utilizando datos demográficos básicos como grupos de edad, sexo y nivel educativo. Para proyectos de investigación existe un procedimiento formal para solicitar acceso controlado a versiones ampliadas.

7. Respetto a la información étnico-racial: ¿por qué hay tanta omisión en el REBP comparado con el censo? ¿Se usa el censo para alimentar esta variable?

Se detectó que alrededor del 67% de la población en el REBP no tiene información étnico-racial, mientras que en el censo esta información estaba mucho más completa. El REBP hasta ahora no utiliza directamente el censo para alimentar esta variable. Sin embargo, a partir del 2023 existe una normativa que obliga a los diferentes registros administrativos a reportar información étnico-racial. Se considera iniciar la inclusión de esta información en futuras versiones del REBP y se discute la posibilidad de usar el censo y censos propios para validar y complementar esta información.

8. ¿Cómo se enfrentan los retos que plantea la cultura y la tradición de ciertos grupos, como pueblos indígenas, respecto al registro de nacimientos y defunciones?

Se reconoce que en estas comunidades el registro puede quedar en la memoria oral y no siempre se formaliza, lo que genera omisión. Se están usando métodos como el dual system con registros administrativos especializados para detectar omisiones, y se trabaja en acuerdos con autoridades tradicionales para mejorar el flujo de información y utilizar censos propios para control y fortalecimiento del registro.

9. ¿Existe un plan o manual para auditoría de calidad y sensibilización para los tomadores de datos, especialmente en salud?

Se reconoce la importancia de desarrollar un manual de auditoría de calidad que considere lo dinámico de los datos. También se ha enfatizado la necesidad de sensibilizar a los profesionales de salud acerca de la importancia de registrar correctamente la información, ya que actualmente su formación no prioriza el manejo adecuado del dato.

10. ¿Qué variables del censo está recogiendo el REBP y cuáles no?

El REBP recoge información a nivel persona y ubicación (dirección, municipio, departamento), pero no tiene estructura de hogar ni vivienda como el censo. En cuanto a variables, incluye datos básicos demográficos, pero no todas las variables sociodemográficas profundas como las que se capturan en el censo, especialmente las étnico-raciales, que se están empezando a incorporar.

11. ¿Cuál es el papel futuro del censo considerando el desarrollo del REBP?

Se espera que a mediano y largo plazo el REBP pueda reemplazar el modelo tradicional del censo, como ocurre en algunos países europeos y Estados Unidos, aunque para ello se

utilizarán datos del censo y censos propios como herramientas de contraste y validación en esta fase inicial.

12. ¿Cómo se manejará la mejora continua de la información y las futuras versiones del REBP?

Se está fortaleciendo la construcción de la versión 2023 del REBP, donde se incorporarán nuevas variables con mejor calidad y cobertura, especialmente la información étnico-racial. También se plantea incorporar análisis de riesgo de reidentificación y mejorar la trazabilidad de los datos, buscando que el REBP permita el análisis de dinámicas demográficas, no solo conteo estático.

5. Compromisos, conclusiones y cierre

La reunión fue un espacio importante de socialización y análisis crítico del Registro Estadístico Base de Población del DANE. Se evidenció el compromiso del equipo técnico y los expertos en fortalecer la calidad, cobertura y utilidad del registro, en particular para reflejar adecuadamente la diversidad étnico-racial del país y avanzar hacia un sistema estadístico más robusto y actualizado.

Se agradeció la participación de los expertos y se reconoció la riqueza de los aportes, especialmente en temas metodológicos y de calidad de datos. Se acordó enviar el acta con las preguntas abordadas durante la sesión y las demás para su retroalimentación posterior. Se programó la próxima sesión para el 25 de julio, enfocada en el Sistema de Información Wayúu. Finalmente, se reiteró el compromiso de continuar fortaleciendo el registro estadístico, base de población y se cerró la sesión con agradecimientos y buenos deseos para los asistentes

Compromisos

Tarea:	Envío de acta para revisión y aprobación
Responsable:	Carlos Eduardo Castro -GIT PAE
Fecha:	02/07/2025
Tarea:	Envío de comentarios escritos preguntas orientadoras
Responsable:	Expertos asesores de la sala
Fecha:	Máximo previo siguiente sesión
Tarea:	Remisión material y asunto de consulta próxima sesión
Responsable:	Juan Sebastian Vásquez y Carlos Eduardo Castro - GIT PAE
Fecha:	10/07/2024

Próxima reunión:

Responsable de convocar: DIRPEN

Fecha: 25/07/2025