**Para el análisis**

El proyecto PRECAV ya tiene un informe que cumple con lo acordado con el PERIS en la concesión de la ayuda. Tras años de espera para obtener los datos de mortalidad estos son limitados y **constituirán una nota menor en los artículos, pero estarán**. El proyecto quería evaluar como los factores de riesgo tenían una distinta prevalencia e impacto según la edad de las personas. En particular interesaba evaluar como en aquellas personas que tenían eventos a una edad prematura la escala tradicional de evaluación de riesgo (REGICOR) tendría menos interés. Tras este tiempo **la cohorte se ha hecho muy antigua (2010-2016)**, corremos el riesgo de no publicarla, y hemos de sacar rédito a estos datos. Ha cambiado también la escala de riesgo que usamos/usaremos.

**Escribiremos SOLO tres artículos.** Uno con toda la población general, uno con DM2, otro con DM1. Describo la idea de cómo organizarlos en general para SACAR TODO EL TRABAJO ADELANTE DE (casi) UNA VEZ Y aprovechar la sintaxis para todos. Cada autor comentará si precisa algún detalle.

1. **POBLACIÓN GENERAL**

La predicción de riesgo y la evaluación del riesgo tiene interés a partir de los 35-40 años. Por este motivo para este estudio se solicitó extracción de datos de personas nacidas en, o más allá de, 1982. La edad mínima será 27 años en 1/1/2010

1. Parte primera epidemiología del primer evento.

**2010:**1) Eliminar personas con un evento previo, 2) eliminar los que tienen DM1 conocida, 3) eliminar los que tienen ACxFA conocida. De esa población total a riesgo identificar a todos aquellos que han tenido un primer evento en el periodo. En ellos hacer la descripción que ya tenemos de eventos (**doc adjunto: 1. Resultats\_poblacionals 2020-10-01**) por territorios, edades, sexos, etc, etc. Son unos 4M de personas y unos 30000 eventos.

De esta población con eventos tenemos una muestra en los que tenemos mortalidad específica. **El análisis de mortalidad con esta muestra con información del INE disponible no cambia (**1. Resultats\_MOSTRA MORTALIDAD 2022-05-12**).** Ya está hecho. Jordi: me tendrás que decir cómo se selecciona la muestra para dejarlo claro en la metodología y decir que es una muestra representativa.

Es una muestra totalmente aleatoria, un 8% (400 000/5 015 212) de la población inicial , y es representativa?. Se puede comparar tabla1 de muestra versus tabla1 poblacional, y se puede comprobar que la edad media y distribución de sexo es idéntica en población que en muestra.

1. Parte segunda: estudio caso control (**ver**1. PRECAV POBLACIONAL CASO CONTROL 2020-03-10)

Una vez hemos hecha la descriptiva por edades y sexo de la primera manifestación cardiovascular, construimos la cohorte de casos (eventos entre 2010 y 2016 de los cuales hemos hecho la descripción) y controles (quedamos en usar 4). Puedes hacer un reporte con muestra para revisar.

Para facilitar el análisis de los otros dos grupos y papers (DM1 y DM2) y no tener que usar otra sintaxis, **mantenemos los CINCO grupos de edad del análisis previo (Y, EA, MA, YO, MVO)**. DEJA en stand by de momento LA SINTAXIS DEL SCORE2 y SCOREOP, si n lo encuentras. Volvemos a ella tras verano SI PODEMOS.

1. **Exclusión.** Eliminar DM1, eventos previos, ACxFA (ya lo hemos hecho antes)
2. Crea casos y controles como lo haces habitualmente, sexo, edad, tiempo de seguimiento, UP, etc
3. Mostrar diagrama de flujo y tabla de distribución de número de controles por cada caso
4. **IDEM. Tabla 1: Características clínicas en fecha de detección/registro población global.** Añade ya en esta tabla las variables específicas creadas para diabetes por Bogdan:
   * MICRO (nefro, neuro y retinopatía**). Aquí se han de identificar los componentes individuales.** Nefro, micro, retino usando el cómo los habéis creado en estudios previos (Bogdan, echa una mano aquí por favor).Crear MICRO\_1 (si tiene cualquiera de las tres), micro\_2 (si tiene 2 de 3), y micro\_3 (si tiene las tres)
   * MACRO (AVC, AP, IC, CI). Por diseño solo estarán en el grupo caso, mantener los grupos grandes en todo caso ya que el descriptivo se ha hecho previamente
   * Duración de DM2 como continua y categorías de 0-5, 5-10, 10-15, > 15
   * Duración de DM1, 0-10, 10-15, 15-20, > 20. BOgdan revisa que estos rangos sean los usados en el otro paper.
5. **IDEM**. Tabla 2: fármacos prescritos activos o dispensados (último año) en fecha de de detección/registro población global
6. Para los análisis de DM habrá que incluir los de la diabetes que ha apuntado Bogdan.
   * FP.NIAD
   * FP.INSULINAS
   * FP.ADO
   * FP.Metformina
   * FP.ISGLT2
   * FP.IDPP4
   * FP.GLP1
   * FP.AGI
   * FP.SU
   * FP.TZD
   * FP.INSULINA RAPIDA
   * FP.INSULINA LENTA
   * FP.INSULINA MIXTA
   * FP.INSULINA INTER.
7. Las tablas del informe 1.2 población joven prematuros y 1.2 Población mayor (no prematuros) pasan a convertirse **en tablas por cada grupo de edad (Y, EA, MA, YO, MVO)**con las mismas variables que la tabla 1 y tabla 2.
8. **Apartado 1.3, 1.3. Por grupos de edad y sexo**. Se puede hacer igual, pero por grupo de edad y sexo. Como seria una tabla muy ancha haz una para total y una cada sexo.
9. SALTAMOS todo lo que tiene que ver con REGICOR. Hablamos cuando encuentres la sintaxis o código R. Debe estar en algún sitio. Te ayudaré a buscarlo.
10. PASAMOS a **“Evaluar componentes del REGICOR en ECV prematuros”.** Lo llamamos **Evaluar peso de FR en eventos según edad. Ver los FR en el Excel. Aquí generar una total, una hombres y una mujeres**
11. Generar las tablas con n y porcentajes de cada uno de los FR que adjunto en el Excel adjunto
    1. Tabla para población general por grupos de edad
    2. Tabla para hombres por grupos de edad
    3. Tabla para mujeres por grupos de edad
12. Luego generar las tablas con los ORy 95% CI (referencia siempre control) para cada grupo de edad.Ya haré yo los forest plot con esa info:
    1. De la población general, (aquí una de las cosas que serán importantes es la comparativa de sexo en cada grupo de edad
    2. De hombres
    3. De mujeres