



EMPIRIA. Revista de Metodología de las Ciencias Sociales

ISSN: 1139-5737

[empiria@poli.uned.es](mailto:empiria@poli.uned.es)

Universidad Nacional de Educación a Distancia  
España

Merino Noé, Jordi

La potencialidad de la Regresión Logística Multinivel. Una propuesta de aplicación en el análisis del estado de salud percibido.

EMPIRIA. Revista de Metodología de las Ciencias Sociales, núm. 36, enero-abril, 2017, pp. 177-211

Universidad Nacional de Educación a Distancia  
Madrid, España

Disponible en: <http://www.redalyc.org/articulo.oa?id=297149614007>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en [redalyc.org](http://redalyc.org)

[redalyc.org](http://redalyc.org)

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

# *La potencialidad de la Regresión Logística Multinivel. Una propuesta de aplicación en el análisis del estado de salud percibido.*

*The potential of Multilevel Logistic Regression. A proposal for implementation in self-perceived-health.*

JORDI MERINO NOÉ

Universitat de Girona  
jordi.merino@udg.edu (ESPAÑA)

**Recibido:** 22.11 2015

**Aceptado:** 08.01.2017

## **RESUMEN**

La inclusión de comandos concretos en aplicativos generales de estimación estadística para analizar modelos multinivel como SPSS, R o STATA está facilitando su extensión en diferentes disciplinas de las ciencias sociales en los últimos años.

Estas técnicas cuantitativas ayudan a observar las características de los individuos y los contextos en los que interaccionan, o niveles en los que son agrupados, en un mismo modelo analítica sobre un variable dependiente.

Aunque inicialmente prosperaron en base a variables dependientes de tipo lineal, también pueden emplearse en modelos con variables de tipo categórico y especialmente también de tipo dicotómica, que es una fórmula muy extendida en sociología y ciencia política a través de la Regresión Logística Multinivel.

El artículo que se presenta a continuación ofrece una simple explicación inicial para familiarizarse con esta técnica y proporciona su aplicación a través de un análisis de caso práctico. Así pues el objetivo es poner de relieve su potencialidad y su posible utilización en investigaciones que deseen incluir factores y covariables contextuales en sus modelos con variables respuesta en función del cumplimiento de un fenómeno de forma factible

El caso práctico escogido responde al desarrollo empírico de diversos debates sobre los macrodeterminantes sociales de salud en el ámbito de la sociología de la salud. Se procura observar la incidencia de variables individuales de las condiciones en las que se trabaja la población asalariada, como es empleada y que exigencias doméstico-familiares tienen, conjuntamente con variables del

mercado de trabajo y de protección social de 27 países de la Unión Europea sobre el estado de salud percibido.

Los principales resultados indican asociaciones ya manifestadas por estudios de salud laboral recientes donde las peores condiciones de trabajo y el empleo temporal se asocian con el mal estado de salud con importantes desigualdades de género entre países. El estudio, pero, ha permitido ir más allá del análisis individual y observar que un porcentaje de las diferencias entre asalariados/as puede ser explicado por las diferencias relativas entre países.

## **PALABRAS CLAVE**

Modelos multinivel categóricos, parámetros aleatorios en modelos multivariados salud laboral, factores de riesgo psicosocial, condiciones de empleo.

## **ABSTRACT**

The development of different software is facilitating the extension of multilevel models in different disciplines of social sciences in recent years. These quantitative techniques help us to observe characteristics of individuals and the contexts in which they interact in the same analytical model. It let to regard the effect of several explanatory variables on a phenomenon considering both the differences between individuals and the available information that they are grouped above in different hierarchical.

Although initially it's thrived in models of linear dependent variables, it could also be used in models with categorical outcomes. The Multilevel logistic regression (MLR) can facilitate estimation of complex models with a dichotomous categorical dependent variable type.

The current article is a brief explanation of this technique. We seek to indicate that it can be used in different areas of sociological analysis, especially in inferential statistics. Despite its apparent complexity, thanks to the recommendations of some authors, we can utilize in a feasible and to facilitate the provision of information on the association of variables and the context of a phenomenon.

In the second part of this text we offer an adjusting exercise of multilevel logistic models. The intention was to observe the incidence of individual variables such as working conditions, occupational conditions and the domestic-family demands that working people have, together with variables of the labour market and social protection of the 27 EU members on perceived health status.

The main results indicated already that the worst working conditions and temporary employment were associated with poor health with significant gender inequalities between countries. These conclusions were noted by some occupational health studies. This study allowed going beyond the individual analysis and observing that a percentage of the differences between employees can be explained by different variables of welfare between countries.

## KEY WORDS

Categorical multilevel models, random parameters in multivariate models, occupational health, psychosocial risk factors, work conditions.

## 1. INTRODUCCIÓN

A pesar que hace ya tres décadas que se comenzaron a utilizar modelos multinivel para analizar datos de población situados o anidados en diferentes niveles jerárquicos de forma conjunta - por ejemplo en el aula, la escuela, el tipo de centro, región...-, ha sido en los últimos años cuando han comenzado a aflorar publicaciones y estudios que los utilizan más allá del campo de la educación, donde siempre han sido bastante presentes (Bliese y Jex, 2002; Heck et al. 2012; Van Yperen y Snijders, 2000). La aparición y evolución de herramientas informáticas han ayudado a que estas técnicas se hayan ido extendiendo en la sociología y las ciencias sociales en general y han ayudado a investigadores e investigadoras a dar respuesta a algunas de sus preguntas de investigación.

Aunque la concepción inicial de los modelos multinivel fuere lineal se ha ido desarrollando también la posibilidad de poder realizar estudios con variables dependientes de tipo categórico. Esto ha permitido que se puedan utilizar modelos donde la variable dependiente esté estructurada de forma dicotómica. Es decir, en función del cumplimiento de un fenómeno o no, fórmula muy extendida no solo en la epidemiología social, y en las ciencias de la salud sino también en la ciencia política. La técnica que se utiliza en estos casos concretos es la desarrollada a partir de los modelos de regresión logística multinivel (MLR en adelante), que supone algunos avances considerables por ejemplo en relación a la regresión logística multivariada (RL)(Manor et al., 2000).

Con los MLR se obtienen unos resultados de un modelo de efectos mixtos que incluye una serie de efectos fijos para las variables o factores predictores y unos efectos aleatorios obtenidos a partir de la información de los niveles en los que se agrupan las unidades de observación o individuos. Desde este punto de vista se puede disponer de información sobre los efectos y un cálculo estimado de la variabilidad de los diferentes niveles e individuos.

En el presente artículo se propone mostrar las potencialidades de este tipo de modelos multinivel de pendientes aleatorias y ayudar a comprender aportación empírica que ofrecen. Por ese motivo se propone un análisis de caso práctico para apreciar su aplicación específica en el ámbito de la sociología de la salud, como es el estudio de la salud percibida en las condiciones de trabajo y empleo de la población asalariada europea.

En una primera parte se introduce a la lógica de la técnica sin más ánimo de familiarizar al lector en los conceptos elementales que faciliten su comprensión y posterior interpretación, sin profundizar a nivel teórico y sin pretender complementar algunos de los manuales existentes.

Así pues el objetivo del siguiente texto es desarrollar un tipo de modelo concreto de pendientes aleatorias de multinivel de una manera que ilustre su uso práctico en una investigación y facilite a los lectores posibles aplicaciones específicas en sus investigaciones

## 2. UNA APROXIMACIÓN A LA LÓGICA MULTINIVEL Y SU UTILIZACIÓN EN CIENCIAS SOCIALES

Una de las muchas preocupaciones empíricas de la sociología, en tanto que disciplina científica, es disponer de un mayor grado de conocimiento de la influencia de la estructura social sobre los individuos de una realidad a estudiar. Los modelos multinivel<sup>1</sup> pretenden y pueden dar algunas respuesta a esta necesidad de analizar la relación entre los sujetos y el contexto donde se desenvuelven o interaccionan a nivel de técnicas cuantitativas (Pardo et al., 2007)

Según Murillo Torrecilla (2008) estos modelos constituyen hasta el momento la metodología de análisis más detallada para el tratamiento de datos jerárquicos, contribuyen a una mejora de la calidad de los resultados y ofrecen la posibilidad de estimar la aportación de cada nivel de análisis, de las interacciones entre las variables individuales y de las variables de su contexto. Desde esta concepción autores como Cebolla Boado (2013) proponen que la lógica que disponen debería ser omnipresente en la investigación social ya que

*“La sociología y la ciencia política, como disciplinas que se encargan de medir y explicar regularidades sociales, hacen una referencia casi constante a la existencia de niveles de análisis agregados, que intervienen por encima del nivel individual, y a variables explicativas que operan por encima de éste. (...) los científicos sociales suelen encontrar explicaciones en distintos niveles cuando identifican regularidades empírica socialmente relevantes” (Cebolla Boado, 2013:13).*

Las mayores contribuciones que se han realizado a la perspectiva de la lógica multinivel en las Ciencias Sociales han sido a partir de estudios de los efectos de escuela y en estudios de comportamiento electoral. Estas aportaciones

---

<sup>1</sup> Existen diferentes denominaciones para referirse a los modelos multinivel. En la literatura sociológica en castellano se acostumbra a manejar modelos lineales multinivel o análisis multinivel. En literatura anglófila se esgrima la denominación *multilevel modelling* como aceptación más extendida. En otras disciplinas y campos de aplicación se utiliza la concepción de modelos jerárquicos colineales, modelos jerárquicos lineales, modelos de efectos de covarianza, modelo de coeficientes al azar y modelos de componentes de varianza (Hox, 2002). Cabe destacar que básicamente estas formas se refieren a los modelos lineales, es decir cuando la variable explicada es de tipo lineal numérica. En el caso que ocupa este artículo referido a modelos que intentan predecir los efectos de una variable de tipo categórica, ya sea binaria o multinominal, o independientemente del método que utilice, se definen como modelos logísticos multinivel, modelos de regresión logística multinivel o incluso modelos generalizados categóricos multinivel. Por coherencia con la literatura sociológica en este artículo se hace referencia a modelos de RL.

han ayudado a explicar la varianza de las variables dependientes y a ganar un considerable grado de solidez a los resultados conociendo la influencia de la dimensión estructural, así como a la fiabilidad de los contrastes estadísticos (Berkhof y Snijders, 2001, p. 133). Cabe destacar que, aparte de la sociología de la educación y la ciencia política, también las ciencias económicas se han servido de esta lógica para interpretar modelos multivariados, aunque solo han utilizado algunos de sus parámetros estadísticos (Heck y Thomas, 2009; Clark y Linzer, 2015). También ha sido en la sociología de la salud que en los últimos años se han incrementado los trabajos que han incidido en los modelos de pendientes aleatorias para explicar los efectos de los contextos sobre los individuos (Muntaner et al., 2011).

En los últimos años el abordaje a través de modelos multinivel ha ido ganando presencia en las ciencias sociales en general y en particular en la sociología. Un ejemplo es el aumento de las publicaciones sobre modelos multinivel en revistas científicas en general<sup>2</sup>. Ahora bien, en revistas exclusivamente sociológicas no hay tantas aportaciones como en otras disciplinas. Además las publicaciones existentes se centran más en la explicación teórica de la técnica para ayudar a entender la lógica tal y como suscribe Pardo et al.(2007). Por tanto -y hasta donde se tiene conocimiento-, existe un menor número de estudios aplicados que estimen modelos multinivel.

Aunque los modelos multinivel se remontan a los años ochenta, no es hasta la aparición de los programas computacionales capaces de resolver las complejas ecuaciones y ofrecer mayor facilidad en la interpretación de los resultados, que su uso se ha extendido del campo de la salud a muchas áreas de las Ciencias Sociales (Catalán-Reyes & Galindo-Villardón, 2003; Heck, Thomas, & Tabata, 2012).

A pesar de las potencialidades de esta lógica y el interés creciente para utilizarla, existen algunas limitaciones que dificultan su extensión y también cierto alejamiento por parte de investigadores e investigadoras. Entre los factores más destacados están la complejidad de las ecuaciones y la amalgama de técnicas dentro de los modelos multinivel. Aparicio y Morera (2007) apuntaban hace años, como gran inconveniente la poca documentación disponible que ayudaba a utilizar las herramientas informáticas de forma específica e interpretar sus resultados. Pero han ido apareciendo recomendaciones y consejos de diferentes autores más recientemente<sup>3</sup> y notables mejoras que facilitan la

---

<sup>2</sup> Existe desde 1989 la revista *Multilevel Modelling Newsletter* que de forma específica se centra básicamente en el desarrollo teórico y en la publicación de investigaciones aplicadas en el ámbito de las Ciencias, de la Estadística, la Medicina y la Biología.

<sup>3</sup> Para superar esta limitación, autores como Ronald H. Heck y Lynn N. Tabata de la Universidad de Hawái y Scoot L. Thomas de la Universidad de Claremont, intentan ofrecer recomendaciones a investigadores/as que desean familiarizarse con estos modelos desde la óptica de la educación y la psicología en sus diferentes publicaciones. Estos consejos que pueden trasladarse a otras disciplinas siendo unos buenos manuales, aunque también existen otros anteriores y que se relacionan en las referencias que ayudan a introducirse en la lógica multinivel (Heck, Thomas, y Tabata, 2012; Heck y Thomas, 2009).

ejecución e interpretación en los diferentes software, a pesar que aún existen diferencias y limitaciones importantes entre ellos<sup>4</sup>.

## 2.1. La estructura jerárquica de datos anidados

En las ciencias sociales y en las ciencias de la salud los datos relativos a los individuos suelen presentarse frecuentemente en diferentes niveles de agrupación o pertenencias. Esto es lo que se denomina estructura jerárquica donde las observaciones están aisladas en un plano individual y posteriormente establecidas en diferentes grupos de niveles superiores de micro a macro. La forma de cómo se organizan y estructuran los grupos ya es una cuestión lógica dependiendo del objeto de estudio, de la pregunta que se formule y de los niveles que se quieran comparar u observar.

Un ejemplo clásico de agrupación de una estructura resultante de un determinada estrategia de muestreo es le disponer información de alumnos y alumnas de un país. Éstos se pueden considerar dentro del grupo aula, también en la escuela, en la región dónde está ubicado el centro educativo... La observación individual que es el alumno/a posteriormente es clasificada en conjuntos de sucesivos niveles. Esta sería una ordenación jerárquica básica, pero existen diferentes tipos de anidación de los individuos.

La forma de agregar a los individuos y grupos a priori es una decisión del investigador/a, y es primordial porque los modelos multinivel trabajan con datos estructurados de forma anidada (Heck et al., 2012).

## 2.2. Evitando la falacia ecológica y la atomística

En los modelos multinivel se suele hablar de los errores que pueden suceder al intentar explicar que las asociaciones estadísticas que se observan en un grupo tienen que sucederse también a nivel individual. Un error básico de la interpretación estadística es suponer que las personas de un grupo tienen las mismas características. Este desacierto es el que se ha denominado en desigualdades sociales de la salud como la falacia ecológica. Las características observadas a nivel global en la población asalariada de un país pueden desaparecer en otro nivel como puede ser en un sector, la empresa o la fórmula contractual. Así pues, no se puede suponer que todos los individuos de un mismo grupo tienen porque tener las mismas características.

Hay otra falacia, que se ha denominado atomística, que corresponde a lo

---

<sup>4</sup> Existen una veintena de programas específicos hasta donde se sabe. Pero el gran avance lo aportan los nuevos comandos de paquetes estadísticos más genéricos como son STATA, SPSS o R que han incorporado procedimientos en sus últimas versiones y han facilitado la aproximación a la RLM. Sin embargo, existen diferencias considerables entre unos y otros ya sea porque no recogen todos los parámetros, en su expresión gráfica, en la configuración o cómo se muestran los resultados.



inverso: las asociaciones que se observan de un fenómeno en el plano individual tienden a generalizarse incorrectamente a nivel general.

Los modelos multinivel pretenden evitar estas falacias a partir de la inclusión de información que se introduce en los modelos con variables de cada uno de los niveles a observar. Desde este punto de vista es un avance en comparación, por ejemplo, con la regresión logística multivariada que tiene en cuenta los individuos pero difícilmente otros niveles o el contexto. Así pues, pueden complementar los modelos de regresión tradicionales porque permiten un modelo para cada nivel de agrupación y ver cómo influyen variables y covariables en cada uno de ellos aportando mayor información y reduciendo el desliz de caer en falacias lógicas (Murillo Torrecilla, 2008; Pérez Fernández, n.d.).

### **2.3. Una clasificación de los modelos multinivel según la variable dependiente**

Como se ha mencionado anteriormente, el análisis multinivel ha permitido avanzar en modelos que incorporan una variable respuesta de tipo categórica, e incluso con agrupaciones de sujetos diferente a la jerárquica, dando un paso más allá de los modelos lineales, que es a partir de los que se desarrollaron.

Pérez Fernández (2012) ofrece en su tesis una forma de clasificar los modelos multinivel en función de la variable respuesta. Básicamente distingue entre los modelos lineales y no lineales. En el primer grupo sitúa modelos con una variable dependiente continua. Reúne así modelos de medias repetidas (que permiten ver las medias de los diferentes niveles), de diseños cruzados (dónde un individuo puede pertenecer a grupos diferentes) y de meta-análisis. En el segundo aúna modelos que contemplan una variable dependiente discreta. Dentro de esta clasificación menciona el Modelo Logístico dónde la variable dependiente se expresa en dos valores de forma dicotómica, el Modelo Multinomial dónde la variable es categórica con más de dos valores, el Modelo Poisson que cuenta las veces o situaciones que se da un suceso y los Modelos de Supervivencia que miden un período de tiempo hasta suceder un episodio predefinido -como padecer una enfermedad, fallecer o encontrar un trabajo -.

## **3. ESTRUCTURA BÁSICA DEL MODELO MULTINIVEL**

En este apartado no se pretende dar una formación extensa sobre regresión logística; sino más bien hacer entender de forma elemental la comprensión de su formulación.

El modelo multinivel clásico parte de la regresión lineal simple dónde se formula un modelo para apreciar la relación de una(s) variable(s) independientes ( $X$ ) sobre otra dependiente ( $Y$ ), teniendo en cuenta un término de aleatoriedad. La ecuación de un modelo de regresión lineal simple pretenden conocer la recta



– que es una buena expresión gráfica- que mejor se ajusta en la disposición de toda la información en una nube de puntos, siendo la ecuación básica de la recta  $Y = \alpha + \beta\chi^5$ . Esta recta ayuda a conocer como una variable puede afectar a otra.

En el lenguaje multinivel se asume que en la relación de  $\chi$  sobre  $Y$  existe un error ( $\varepsilon_i$ ) ya que se considera la no independencia de las observaciones. Así pues, la premisa básica del modelo lineal clásico se incumple, porque “*se viola uno de los principios básicos de la estadística frecuentista, la independencia de los datos*” (Pérez Fernández, n.d., p. 11).

Además, “*las hipótesis de partida de estos modelos nos dicen que los individuos pertenecientes a un mismo contexto tenderán a ser más similares en su comportamiento entre sí, que respecto a su pertenencia a distintos contextos*” (Andréu Abela, 2011, p. 163). Se entendería, por ejemplo, que las condiciones de trabajo de los asalariados de una empresa se asemejan más entre ellos que con los de otras empresas, o el comportamiento electoral o hábitos de consumo de los residentes en un barrio en comparación a otros de la misma ciudad. Se plantea por tanto el modelo multinivel básico a partir de la ecuación siguiente:

$$Y_i = \beta_0 + \beta_1\chi_{1i} + \varepsilon_i. \quad (1)$$

A partir de esta ecuación básica se va complementando y haciendo más compleja a medida que se avanza en los modelos, según el método a utilizar, en función de los niveles a considerar, del número de variables que se utilizan y del tipo de variable dependiente (Cebolla Boado, 2013).

### 3.1. La regresión logística multinivel (MLR)

El presente artículo se centra en la variable respuesta de tipo categórica en dos valores, por tanto el modelo logístico. Para entender la MLR debemos partir del objetivo que plantea una regresión logística que intenta predecir una probabilidad que ocurra un fenómeno o acontecimiento según la información que ofrecen las diferentes observaciones. De este modo se puede entrever la probabilidad ( $\pi_i$ )- dentro de un intervalo de confianza – para que ocurra un suceso ( $Y$ ), mediatizado por un grupo de variables independientes o factores  $\chi_1, \chi_2, \dots, \chi_n$  en un individuo ( $i$ ). Entonces la ecuación para una variable dependiente dicotómica busca conocer la probabilidad que suceda  $\pi_i$  y su expresión es:

<sup>5</sup> La pretensión del análisis de regresión lineal simple es poder explicar el comportamiento de una variable por los valores de otra(s). En la ecuación  $\alpha$  es el valor que tiene la variable  $Y$  cuando  $\chi$  es igual a cero. Es decir cuando no hay relación, siendo el punto de corte de la recta. Por otro lado  $\beta$  determina la pendiente de la recta.

$$f(\pi_i) = \beta_0 + \beta_1 \chi_i^6 \quad (2)^6$$

Ésta sería la expresión para el nivel de los individuos, o primer nivel. La intención de los modelos multinivel es ir agrupando los individuos/observaciones en niveles superiores ordenados jerárquicamente ( $j$ ).

En relación a los modelos dicotómicos, la variable dependiente se expresa en el valor 1 cuando sucede el evento y 0 cuando no es así. Entonces, la intención es averiguar la probabilidad que suceda  $Y_{ij} = 1$  en función del individuo ( $i$ ) y el nivel de agrupación ( $j$ ).

Desde esta formulación Pérez Fernández (2012) expresa en el caso de un modelo nulo, que solo contempla los datos de los individuos y sus agrupaciones para una variable dicotómica dependiente el modelo, la ecuación de la siguiente manera:

$$\text{logit}(\pi_i) = \log \text{odds} = \log \left( \frac{\pi_i}{1 - \pi_i} \right) = \beta_{oj} \quad (4)$$

En un modelo con una solo factor independiente para dos niveles la intención es conocer la  $\beta_{oj}$ , que ofrece un efecto fijo del intercepto con  $X$  ( $\beta_o$ ) igual para todos los grupos anidados y un componente de efecto aleatorio específico de cada grupo ( $u_{oj}$ ).

$$\text{logit}(\pi_i) = \beta_o + \beta_1 \chi_{(1)ij} + u_{oj} \quad (5)$$

Esta es pues la expresión algebraica más escueta para entender un modelo básico de MLR. A partir de esta se va haciendo más compleja en función de los diferentes niveles y variables. Las herramientas de los diferentes paquetes estadísticos como R, STATA y SPSS analizan con comandos estas ecuaciones.

### 3.2. Los parámetros fijos y aleatorios

El análisis multinivel permite estimar la contribución de variables independientes situadas en distintos niveles sobre un fenómeno. Además proporcionan una bondad de ajuste de los diferentes modelos y qué porcentaje de la varianza se explica por cada uno de los niveles en cada uno de ellos. A la

<sup>6</sup> En la ecuación 2 aparece una función de probabilidad del fenómeno  $f(\pi_i)$ , siendo una transformación denominada link. Para el caso de la MLR se suele utilizar el link logístico que se expresa de la siguiente manera:

$$f(\pi_i) = \log \left( \frac{\pi_i}{1 - \pi_i} \right) \quad (3)$$

hora de exponer y analizar los resultados, sobre todo, hay dos grandes tipos de parámetros - también denominados coeficientes- que son los de efectos de tipo fijo y los de efectos aleatorios. Estos parámetros facilitan un marco estadístico pertinente para estructuras de datos que presentan diferentes niveles.

Para entender los parámetros de efectos fijos se puede hacer referencia a los modelos de regresión clásicos, dónde se centra la atención en el punto de corte y en la pendiente de la recta. Estos parámetros se entienden o asumen como fijos, ya que son comunes para todos los individuos de la muestra de estudio, y los resultados son una estimación a partir de los datos que éstos ofrecen. En ese sentido tienen en cuenta todos los individuos estudiados, como podría ser por ejemplo todos los trabajadores de una empresa o los alumnos de un aula.

En el tipo de modelo que no ocupa, la MLR los efectos fijos pueden interpretarse de diferentes maneras. Una primera interpretación es la expresión del coeficiente de la intersección o intercepto. Los diferentes procesadores estadísticos suelen ofrecer este parámetro juntamente con su desviación estándar y su nivel de significación.

Una segunda es a través del cálculo de la probabilidad utilizando la función logística inversa. Es decir dividir 1 por 1 más el exponente negativo del intercepto. Esto nos ofrece una probabilidad que si se multiplica por cien se puede interpretar en tanto que porcentaje.

Una tercera, y es la que autores como Heck et al. (2012:156) aconsejan en la regresión logística siendo la más extendida en sociología de la salud - y que es en la que se centrará el ejercicio posterior-, es a partir de la interpretación exponencial del intercepto para observar la influencia de cada variable. Este exponente pasa a denominarse oportunidad relativa asociada (ORA)<sup>7</sup>, que mediante un intervalo de confianza facilita una proporción de cómo determina una variable un fenómeno/evento comparando las observaciones en las que acontece y en las que no. La ORA es muy extendida en Ciencias de la Salud tanto porque permite identificar qué factores pueden ser protectores ante una enfermedad o cuáles son factores de riesgo, así como ayudar a predecir la probabilidad que el evento suceda (Jovell, 2006:29).

Estos parámetros expresados en ORA ofrecen un ejercicio de comparabilidad con otras técnicas como es la RL<sup>8</sup> - tan extensa en la sociología de la salud y la epidemiología social-, ya que su expresión es muy similar (Manor, et al., 2000; Pérez Fernández, n.d.). Pero además, permiten en la MLR observar como varían las asociaciones y la incidencia sobre la variable dependiente en los diferentes modelos o agrupaciones de variables. Las ORA

<sup>7</sup> En la literatura castellana la traducción de odds ratio se ha interpretado de diferentes formas. En este texto se utiliza oportunidad relativa asociada, que junto a razón del riesgo relativo son las más extendidas.

<sup>8</sup> En la tabla 10 se ofrece una comparación de las ORA del modelo utilizado en el caso práctico con las de del mismo modelo utilizando una regresión logística binaria. Como se aprecia en las características individuales las ORA varían poco y solo en algún factor pasa a ser significativa. En las variables de niveles superiores varía en función si es una covariable lineal o una variable categórica. En estos casos siempre es más recomendable utilizar una RLM.

de los parámetros fijos pueden ser tanto de factores de un primer nivel como de variables o covariables del segundo nivel o superiores. Ahora bien, hay que tener en cuenta que éstos congelan las diferencias entre los grupos y no facilitan ningún pronunciamiento sobre qué parte de la varianza puede ser explicada por la diferencia entre los grupos del segundo nivel. Así pues son resultados comunes para todos los contextos. Observar estos efectos o parámetros fijos es ya un ejercicio de regresión multinivel.(Cebolla Boado, 2013).

**Tabla 1. Resultados de los coeficientes fijos de las condiciones de empleo sobre el mal estado de salud y agrupados los individuos según países.**

Término del modelo	Coeficiente	error estándar	t	Significación	Exponente del coeficiente	Intervalo confianza del 95% para el Exponente del coeficiente	
						Inferior	Superior
Intercepto	-3,423	0,158	-21,628	0,000	0,03	0,02	0,04
Sin contrato	0,395	0,104	3,811	0,000	1,48	1,21	1,81

Fuente: Elaboración propia a partir de los datos de la EWCS 2010

Para ayudar a ver mejor lo explicado se facilita un ejemplo recogido del estudio de caso posterior, donde se ofrecen los resultados de los coeficientes fijos para un modelo que considera el mal estado de salud de población femenina asalariada de la UE según países con determinadas condiciones de empleo. Si se desea ver la asociación de trabajar sin contrato sobre el estado de salud percibido, ajustada por edad, el resultado del coeficiente es de 0,395 siendo significativo. El exponente de este valor es de 1,48 ( $2.7182^{0.395}=1,48$ ). Por tanto describe una ORA de 1,48 en un intervalo de confianza del 95% de 1,21 a 1,81. Se podría leer entonces que hay una asociación estadísticamente significativa entre trabajar sin contrato y tener mal estado de salud en mujeres de la UE. Además, se podría expresar que por cada individuo sin contrato con buena salud se observarían 1,48 con mal estado, o lo que es lo mismo que una persona sin contrato tiene esa proporción de expresar una salud no buena.

Esto sería refiriéndo referente a los parámetros fijos, pero otra gran aportación de los modelos multinivel son los parámetros aleatorios dónde son considerados como variables y su distribución es a partir de una función de probabilidad (Murillo, 2008). A partir de la expresión de éstos se puede tener una mejor descripción y especificidad de la diferencia entre los grupos y los diferentes niveles (Merlo et al., 2006).

Existen importantes diferencias en los modelos multinivel en los parámetros aleatorios. Por ejemplo en los modelos lineales multinivel los parámetros fijos analizan las medias de la población, las pendientes y la intersección de una recta, mientras que en los modelos aleatorios el interés se centra en las varianzas y

covarianzas de todos los niveles considerados, en tanto que parámetros asociados a efectos eventuales. Con esta información se puede calcular el coeficiente de correlación intraclase (ICC) que estima qué porcentaje de la varianza del modelo puede ser explicada por la desigualdad entre grupos e individuos. En el caso de los modelos de MLR los parámetros aleatorios son discordantes. Recuérdese que la variable respuesta solo tiene dos valores y al ser binomial no existe varianza residual.

A pesar de esta limitación, se pueden disponer de unos coeficientes estimativos bastante explicativos para interpretar la varianza gracias a las aportaciones de diferentes autores, y que se describen en el siguiente apartado.

#### 4. LA EXPLICACIÓN DE LA VARIANZA EN LOS MODELOS LOGÍSTICOS MULTINIVEL

Al no disponer de varianza residual en la MLR no se puede calcular el ICC que es el coeficiente que explica qué porcentaje de varianza explicada se debe a las diferencias entre los grupos básicamente.

Merlo et al. (Merlo et al., 2006) proponen superar este inconveniente a través de un cálculo que permite estimarlo en modelos categóricos<sup>9</sup>. Aconsejan además otro coeficiente que también se acostumbra a utilizar en algunos estudios de epidemiología social que es la mediana de la oportunidad relativa (MOR, *median odds ratio*). Ambos coeficientes se detallan a continuación.

##### 4.1. El Coeficiente de Correlación Intraclase (ICC)

También denominado como coeficiente de partición de la varianza (VPC) en la literatura inglesa, y teniendo en cuenta las limitaciones expuestas, se puede intentar calcular a través de una ecuación que propone el método conocido como modelo de la variable latente (*threshold model method*). Se plantea desde esta lógica considerar una probabilidad a través de una constante para poder disponer de la información sobre la varianza explicada. Es decir, considerar la varianza individual como una escala de probabilidad igual a  $\pi^2/3$ , o lo que es lo mismo como una constante de 3,29. Incidir que también existe otro método de aproximación al ICC conocido como modelo de simulación residual dónde a diferencia del anterior se suma directamente la constante de  $\pi$ .

*“These methods for computing the ICC in logistic models have their own statistical consistency. However, they are an attempt to apply to the logistic case notions that are based on the clear distinction between the individual level variance and the area level variance that exists in the linear case. As this distinction is not so clear in the logistic case, the interpretation of the*

<sup>9</sup> Siempre que no exista alguna variable en el modelo que denote grandes diferencias entre los grupos y que pueda comportar algún problema en la resolución de la ecuación

*ICC for dichotomous outcomes is difficult to understand in epidemiological terms (Merlo et al., 2006:292)“.*

Con esta consideración y siguiendo las indicaciones del método de la variable latente, se puede intentar calcular el ICC, conocido también como Rho ( $\rho$ ) a través de una ecuación similar a la de los modelos lineales pero con la diferencia de trasladarla a una escala logística:

$$\rho = \frac{\sigma^2}{\sigma^2 + \pi^2/3} \quad (6)$$

Así pues el ICC puede ser entendido como la proporción de la varianza que se explica por las diferencias entre los grupos y los individuos, ya que es una media de la semejanza de los individuos en el grupo del nivel inferior y la diferencia en el nivel más macro o global.

Para entenderlo gráficamente, en la *Tabla 6* se ofrecen la estimación del efecto aleatorio de un modelo nulo donde se tienen en cuenta países y mal estado de salud. La varianza del contexto o nivel resulta de 0,145, por tanto su ICC será de 4,22% que se puede leer como parte de la varianza del mal estado de salud que se puede explicar por las diferencias entre países.

$$\rho = \frac{\sigma^2}{\sigma^2 + \pi^2/3} \quad / \quad \rho = \frac{0,145}{0,145 + \pi^2/3} \quad / \quad \rho = 0,0422 * 100 \quad (7)$$

## 4.2. La mediana de la oportunidad relativa (MOR)

Este otro parámetro puede complementar la necesidad de explicar la varianza del segundo nivel o superiores en los modelos MLR. Hasta donde se conoce, no es un parámetro que se acostumbre a utilizar en la sociología ni en la ciencia política siendo más habitual encontrarlo en estudios de bioestadística y epidemiología. Diferentes autores de estas disciplinas aconsejan su utilización porque ofrece una interpretación de las diferencias de manera más gráfica y entendible, así como una mejor comprensión de las diferencias entre grupos en cada modelo de forma secuencial.

Este indicador básicamente lo que hace es cuantificar qué proporción de la varianza entre grupos existe comparando un individuo de un grupo con el mayor riesgo con respecto a otro grupo dónde se observa el menor riesgo o efecto medido en relación a la variable dependiente. Es decir, acentúa las diferencias entre grupos ya que es una comparación entre dos individuos con las mismas características. En términos epidemiológicos se podría decir que interpretaría el aumento del riesgo mediano que un individuo tendría si cambiara su grupo de referencia por otro.

La interpretación del resultado es similar a los parámetros fijos de la ORA que muestran una proporción, pero sin existir un intervalo de confianza. En este sentido si la magnitud de la varianza que muestra la MOR es un valor alto indicará que existen diferencias considerables entre los niveles de agrupación. Un valor de 1 significaría que no hay variación entre los grupos, siendo valores por encima de 1 mayores diferencias.

Para calcular la MOR, Merlo et al.(2006) proponen una ecuación que depende del área del nivel la varianza. Para ello se multiplica la varianza del nivel por dos y posteriormente por el setenta y cinco percentil acumulado en una distribución normal con media 0 y varianza 1. Proponen su medida con la ecuación siguiente para cada modelo que se realice, considerando que es estadísticamente independiente de la prevalencia del fenómeno estudiado<sup>10</sup>:

$$MOR = \exp(\sqrt{2 \times \sigma^2} \times \phi^{-1}(0,75)) \quad (8)$$

## 5. EL PROCESO DE ESTIMACIÓN Y MODELIZACIÓN

Al igual que otras técnicas estadísticas, en el proceso de estimación lo que se busca es ir construyendo un modelo en función de los objetivos, las preguntas que se formulan y la disponibilidad de información. Consiste así, en encontrar un modelo lo más acertado que permita entender el efecto de las diferentes variables independientes o predictivas a medida que se van incluyendo sobre la variable independiente o fenómeno.

El proceso de estimación ayuda a encontrar un buen modelo final, pero sobre todo ampara a observar el efecto y relación de las diferentes agrupaciones de variables sobre el suceso en un modelo multivariado. En los modelos lineales es habitual en este ejercicio combinar variables de primer nivel y superiores para ver como la inserción de una variable controlada puede explicar una parte de la varianza o las disimilitudes observadas entre los diferentes grupos. De este modo se puede apreciar si los parámetros de bondad mejoran, si la nueva variable/es explica/n parte de la varianza o si la variabilidad entre los grupos en el segundo nivel o superiores se ve alterada en los efectos aleatorios.

En los modelos logísticos no suele ser así: a pesar de que se puede disponer del coeficiente de cuasiverosimilitud y del coeficiente de *Aikake* utilizados en los modelos lineales, se ha de proceder con cautela al utilizarlos en la RLM ya que a menudo la transformación de los datos es diferente. Por este motivo diferentes autores relativizan la interpretación de estos índices y no se acostumbran a emplear.

<sup>10</sup> Algunos procesadores no facilitan el cálculo de las ecuaciones del ICC ni la MOR. Se aconseja realizarlos mediante un procesador de hojas de cálculo a través de la  $\sigma^2$  que ofrecen las ventanas de resultados.



El primero de los modelos en este proceso acostumbra a ser uno que incluye únicamente la variable a explicar. Esta variable es estimada como una constante más un efecto aleatorio en función de los individuos y los grupos en los que éstos se han organizado para el análisis. La importancia de este primer modelo es ofrecer información sobre si es pertinente o no continuar con el proceso de estimación y ver si hay diferencias entre los niveles de agrupación sin incorporar ninguna variable o covariable que pueda explicar la varianza. Es a partir de la información de este primer modelo que se establece la estrategia metodológica de ir incorporando de forma sucesiva el resto de variables y covariables definidas previamente para intentar reducir la varianza no explicada y ver así los posibles efectos sobre el fenómeno (Pérez Fernández, n.d.).

Este primer modelo recibe el nombre de modelo vacío o modelo nulo (en la literatura inglesa *null model* o *empty model*), aunque también se puede encontrar en la literatura como modelo incondicional de medias. Este es el punto de partida y la base del proceso al ser el modelo más simple de descomposición de la varianza. Ofrece la intercepción o constante, por tanto facilita pesquisas sobre la varianza de la variable dependiente sin ninguna variable predictiva y permite ver cómo evolucionan los diferentes niveles en los siguientes modelos.

A partir del modelo nulo, se van incorporando las agrupaciones de variables y se puede ir observando los parámetros fijos y aleatorios de cada modelo, hasta llegar a un modelo lo más acertado para las preguntas planteadas.

No existe una regla o norma de cómo ha de ser el proceso de estimación o cual es el modelo más acertado. Sí que concurren, sin embargo, diversas recomendaciones en el proceso para evitar caer en una estrategia de minería de datos – o de *fishing expedition* – sin tener una idea clara de qué variables introducir en el modelo. Los principales consejos que ofrecen diferentes autores son que se tenga en cuenta el principio de parsimonia – que asevera en explicar un fenómeno o problema la forma más sencilla y explícita como la más apropiada–, que los modelos utilizados se sustenten en modelos teóricos y sobre todo que lo hagan en función de la hipótesis que se desea contrastar (Balluerka, Gorostiaga, Gómez-Benito, & Hidalgo, 2010; Cebolla Boado, 2013; Merlo et al., 2006; O’Connell & McCoach, 2008; Pardo, Ángel, & San, 2007; Pérez Fernández, n.d.).

## **6. CASO PRÁCTICO: EL ANÁLISIS DEL ESTADO DE SALUD PERCIBIDO EN EUROPA.**

En el debate sobre macrodeterminantes sociales de la salud a nivel europeo ha emergido en la última una década una atención centrada en la incidencia de los Regímenes de Bienestar, en las condiciones de cómo la población es empleada, en qué condiciones realiza el trabajo y qué relación hay con el ámbito del trabajo reproductivo. Diferentes autores han suscrito la necesidad de emplear modelos que visualicen a nivel empírico lo descrito a nivel teórico donde se interrelacionen las características individuales con elementos de

economía política de la salud y de tipo de políticas públicas de cada país o región que faciliten la comprensión de los efectos sobre la salud, las diferencias entre territorios de la UE y el porqué del incremento o mantenimiento de las desigualdades sociales en salud (Bambra y Eikemo, 2009; Benach y Tarafa, 2010; Benach et al., 2014; Muntaner et al., 2011)..

Además, en los últimos años han aumentado las investigaciones que han recogido estas sugerencias y han comenzado a analizar de forma conjunta las condiciones de empleo y trabajo, aunque se han centrado en alguna de las dimensiones o en estudios comparativos entre países o tipos de Regímenes (Bambra et al., 2014; Bliese y Jex, 2002; Campos-Serna et al., 2013; Huijts, Perkins, y Subramanian, 2010)

A continuación se presenta un modelo que intenta recoger los principales ámbitos de interés en este debate. El caso práctico pretende describir la influencia que tienen sobre el estado de salud percibido, diversas variables individuales sociodemográficas, de condiciones de empleo, de trabajo, de las exigencias doméstico-familiares y de covariables contextuales del mercado de trabajo y del sistema de protección social de la población asalariada de 27 países de la UE.

El estudio se basa en dos hipótesis: La primera es que existen asociaciones estadísticamente significativas tanto de las variables individuales como de contexto con el estado de salud percibido de la población asalariada de la UE, siendo estas asociaciones observables en un modelo multivariado donde se consideran diferentes dimensiones del trabajo productivo y reproductivo y de variables comparativas de Regímenes de Bienestar. La segunda es que una parte de las divergencias entre los individuos es debido a las diferencias entre los países.

## 6.1 Métodos y proceso.

**Población de estudio:** La población de estudio fueron los 16.891 asalariados y las 14.948 asalariadas de los 27 países miembros de la UE en el año 2010 que contestaron la encuesta de la quinta edición de la *European Working Conditions* (EWCS) elaborada por la Fundación Europea de Condiciones de Trabajo y Salud (EUROFOUND). La distribución por países se facilita en la tabla 8 y 9 al final del documento.

**Variables:** El conjunto de variables incorporadas en el análisis, tanto de carácter individual como contextual, se recogen en la Tabla 2.

**Tabla 2. Relación de las variables incorporadas en el análisis según nivel jerárquico.**

Tipo de variables		Variables medidas en el estudio
variable dependiente		Estado de Salud
Variables independientes	nivel jerárquico 1	Sociodemográficas
		Edad
		Sexo
		Mercado de Trabajo
		Tipo de Contrato (fijo, temporal, vía ETT, aprendizaje, sin contrato)
		Inseguridad en el trabajo
		Tiempo parcial
		Entorno Psicosocial del trabajo
		Ex. Psicológicas (cuantitativas y cognitivas)
		Control del trabajo (previsibilidad, control)
		Apoyo social (superiores y compañeros/as)
		Recompensas del Trabajo
		Sentido del Trabajo
		Exigencias doméstico-familiares
		Ex. Atención y educación hijos e hijas
		Ex. Atención hogar y cocinar
		Ex. Gente mayor y dependientes
	nivel jerárquico 2	Mercado de Trabajo
		Tasa paro
		Paro larga duración
		Variación tasa paro 2005-2010
		Protección social/pobreza
		Gasto en protección social en upc/habitante 2010
		Población asalariada por debajo umbral de pobreza

Fuente: Modelo de elaboración propia a tenor de diferentes debates en el ámbito de la salud laboral y desigualdades sociales de la salud relativos a los efectos macro en la salud percibida de las condiciones de empleo en población asalariada.

La variable dependiente fue el estado de salud percibido que se recogía directamente en el cuestionario y que estaba distribuida en cinco valores: muy bueno, bueno, pasable (o regular según traducción), malo y muy malo. Al tener una distribución categórica se dicotomizó entre estado de salud muy bueno y bueno versus estado de salud no bueno siguiendo el criterio utilizado en la gran mayoría de estudios epidemiológicos que utilizan esta variable como fenómeno a

explicar en comparación entre países (C Bamba y Eikemo, 2009; Bamba et al., 2014; Campos-Serna et al., 2013; Eikemo, Bamba et al., 2008; Huijts, Perkins, y Subramanian, 2010).

Las variables individuales se concentraron en condiciones de ocupación, condiciones de trabajo y exigencias doméstico-familiares. Las variables de las condiciones de ocupación fueron el tipo de contrato (fijo, temporal, vía ETT, aprendizaje o prácticas y trabajar sin contrato), el trabajo a tiempo parcial (jornadas semanales de menos de 35 horas para homogeneizar una media en la UE) y la inseguridad laboral que consideró solo el valor extremo de estar muy de acuerdo en perder el trabajo en los próximos seis meses.

Para las variables condiciones de trabajo se consideraron únicamente los factores de riesgo del entorno psicosocial del trabajo no teniendo en cuenta otros factores de tipo físico, químico o ergonómico. Para medir estos factores se construyeron las variables a partir de las indicaciones de los modelos teóricos demanda-control-apoyo social y el de desequilibrios entre esfuerzos y recompensas, y siguiendo las indicaciones del cuestionario COPSOQ-ISTAS 21 (Karasek & Theorell, 1999; Salanova Soria & Schaufeli, 2002; Siegrist, 2011). Se elaboraron diferentes escalas factoriales siguiendo las orientaciones de Moncada et al., (2014) y a partir de la comprobación de la consistencia interna, se creó una nueva variable para cada factor que consideró la media de los diferentes ítems a partir de la mediana al no ofrecer una distribución normal. La consistencia interna que fue aceptable en todos los casos.

Las exigencias doméstico-familiares se construyeron a partir de reconocer en el cuestionario la realización o no de tareas de atención y educación de los hijos e hijas, tareas de atención del hogar y cocinar, y tareas de atención y cuidado a gente mayor o dependiente. No se pudieron hacer opciones de recuentos de horarios ni otras medidas de encuestas de usos del tiempo porque el cuestionario no ofrecía estas posibilidades.

Las variables contextuales del segundo nivel se crearon a partir de los indicadores que ofrecía el Servicio de Estadística de la UE EUROSTAT para el año 2010 según sexo<sup>11</sup>, y se agruparon entre las que aportaban información sobre el mercado de trabajo y las que aportaban información sobre la protección social y la pobreza de la población asalariada.

Para el primer grupo se consideró la tasa de paro en el último trimestre del año 2010, el paro de larga duración y la variación de la tasa del país del año 2005 al 2010. Para el segundo grupo se consideró el gasto social en protección social estandarizado en unidades de poder de compra por habitante, y el porcentaje de población asalariada por debajo del umbral de pobreza. El gasto en protección social estandarizado se categorizó en países con menos de 3.000,00, de 3.000,01 a 6.000,00, de 6.000,01 a 9.000,00 y más de 9.000,01 upc/habitante. La otra variable se categorizó en porcentaje de población asalariada por debajo

<sup>11</sup> Probablemente se podrían haber incorporado otras variables para visualizar las diferencias entre los países de la UE, pero no todos los indicadores se ofrecen de forma sexuada para los 27 países.

del 5%, de 5,01% a 10,00% y más de 10,01%.

Como variables demográficas se consideraron la edad y el sexo que se recogían directamente en el cuestionario. El análisis descriptivo de las variables del estudio se recoge en la *Tabla 9*.

**Proceso:** Después de ponderar la muestra en función de los criterios territoriales, demográficos y del mercado de trabajo, los datos se analizaron a través de la realización de modelos de regresión logística multinivel. El software utilizado fue el paquete estadístico *IBM SPSS statistics 19.0* a través del procedimiento GLMM o modelos mixtos categóricos con el método de máxima log pseudoverosimilitud o de cuasiverosimilitud (PQL)<sup>12</sup>. Para la realización de las ecuaciones de los parámetros aleatorios del ICC y la MOR se complementó con procesadores de cálculo ya que no son ofrecidos por el procesador estadístico, y siguiendo las propuesta de Merlo et al. (2010) y las orientaciones de Heck et al (2012) y de Rabe-Hesketh y Skrondal (2006).

El proceso secuencial de estimación constó de 6 modelos realizados de forma independiente para hombres y para mujeres. La estrategia de modelización se facilita en la *Tabla 3*<sup>13</sup>.

---

<sup>12</sup> El método de máxima verosimilitud es una técnica que ayuda a confeccionar los parámetros a partir de la estimación de los valores de una muestra finita de datos. En el caso de los métodos de cuasiverosimilitud se permite una mayor variabilidad de los datos y es la que se aconseja a por Heck et al. (2012) para distribuciones binomiales en modelos mixtos.

<sup>13</sup> En este artículo se incide en la importancia del proceso secuencial para facilitar la interpretación y familiarizarse con modelos logísticos multivariados, sobre todo a investigadores noveles. Habitualmente se suele centrar la atención en el modelo final y se muestra también como interesante el modelo nulo o vacío.

**Tabla 3. Estrategia de modelización del proceso de estimación en el análisis logístico multinivel.**

Nivel jerárquico	Co-variables	Modelo nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
1	Edad		*	*	*	*	*
	Tipo de contrato		*	*	*	*	*
	Contrato a tiempo parcial		*	*	*	*	*
	Inseguridad laboral		*	*	*	*	*
	Ex. Psicológicas cuantitativas			*	*	*	*
	Ex. Psicológicas cognitivas			*	*	*	*
	Previsibilidad			*	*	*	*
	Control			*	*	*	*
	Apoyo Social			*	*	*	*
	Sentimiento del trabajo			*	*	*	*
	Recompensas			*	*	*	*
	Ex. Atención Hijos/as				*	*	*
	Ex. Mantenimiento hogar y cocinar				*	*	*
	Ex. Atención gente mayor y/o dependientes				*	*	*
2	Tasa paro IV2010					*	*
	Paro larga duración 2010					*	*
	Variación tasa paro 2005-2010					*	*
	Gasto en protección social 2010						*
	Población asalariada bajo el umbral de pobreza						*

## 6.2. Resultados

### 6.2.1. El ejemplo del modelo nulo

Con la información que aporta el modelo nulo y no teniendo en cuenta el resto de características individuales, se observa en la tabla 4 que el exponente de  $\beta$ , o ORA era de 0,24 en hombres y de 0,26 en mujeres estando dentro del intervalo de confianza del 95%. En ese sentido existía una relación o asociación positiva entre el estado de salud y las diferencias entre países. Otra interpretación era que la probabilidad de tener mal estado de salud según las diferencias entre países

era de 19,43% en los hombres y de 20,66% para las mujeres, sin considerar el resto de variables<sup>14</sup>.

No se puede aportar demasiada información sobre la bondad del ajuste del modelo, o si es interesante continuar observando qué variables pueden incidir en el fenómeno a partir del proceso secuencial, ya que la MLR no dispone de otros indicadores que ayuden a esta interpretación.

### 6.2.2. Modelos secuenciales y parámetros fijos

Los resultados de los parámetros fijos de los modelos de regresión logística multinivel, mediante las oportunidades relativas asociadas (ORA) con un intervalo de confianza del 95% quedan recogidas en la Tabla 4 y la Tabla 5.

El primer modelo contemplaba la edad y las condiciones de ocupación. La edad fue estadísticamente significativa tanto en el primer como en posteriores modelos, y se asoció de forma positiva por encima del valor 1, *per ende* influyendo sobre la variable explicada en ambos sexos.

Las condiciones de ocupación en todos los modelos se asociaron con el estado de salud percibido. Los asalariados y asalariadas con un contrato temporal fueron más propensos a contestar que no tenían buena salud. En modelo final la ORA fue de 1,25 (IC95% 1,07-1,47) en hombres y de 1,30 (IC95% 1,11-1,51) en mujeres. El mismo efecto se apreció con las variables de precariedad laboral siendo estadísticamente significativo para los dos sexos en todos los modelos. En el caso de los asalariados contratados en la fórmula de aprendizaje o prácticas se denotó una ORA elevada. En las asalariadas trabajar sin contrato era un factor de riesgo, fenómeno que no se apreció en los hombres. Por otro lado el contrato a tiempo parcial se comportó como un efecto protector de la salud sólo en hombres (ORA= 0,84 IC95% 0,71-0,99). Esto se podría interpretar también como qué trabajar a tiempo parcial sería un 16% más protector del mal estado de salud que hacerlo a tiempo completo. El análisis combinado de las estimaciones fijas no hizo variar la dirección de estas variables con la incorporación de otros grupos de variables o covariables de segundo nivel en modelos posteriores.

Los resultados de los parámetros fijos de las condiciones de trabajo-incorporadas a partir del segundo modelo-, fueron más diversos y en casi todas las variables se observaron resultados estadísticamente significativos. Se comportaron en la misma dirección que indicaban los modelos teóricos de los factores de riesgo psicosocial para ambos sexos. Las exigencias psicológicas altas, la poca previsibilidad y las bajas recompensas por el trabajo se comportaron como factores de riesgo. Por otro lado el control sobre el contenido del trabajo, el apoyo social y el sentimiento del trabajo se asociaron de forma

<sup>14</sup> Esta proporción surge a raíz de transformar el exponente de  $\beta$  en una probabilidad que el fenómeno suceda en una función logística a partir de la ecuación:

$$Y = \frac{\exp(\beta)}{1 + \exp(\beta)} \times 100 \quad / \quad Y = \frac{0,24}{1 + 0,24} \times 100 = 19,43$$



negativa indicando su efecto protector sobre la salud.

Las asociaciones con el mal estado de salud relacionadas con las exigencias domésticas familiares se incorporaron a partir del modelo 3. Éstas no fueron significativas en ningún modelo para los hombres mientras que en el caso de las mujeres lo fueron las exigencias de atención y educación a los hijos e hijas (OR= 0,89 IC95% 0,80-1,00), y las exigencias de atención y cuidado a gente mayor y/o dependiente (OR= 1,44 IC95% 1,17-1,76).

Las variables contextuales del segundo nivel se incorporaron a partir del modelo 4 y aportaron escasas asociaciones estadísticamente significativas. Como se ha mencionado, su consideración en los modelos no hizo cambiar ni la dirección de las asociaciones de los grupos de variables anteriores ni demasiado las ORA.

En el modelo 4 en hombres se observó, dentro de las variables del mercado de trabajo, que la variación de la tasa de paro entre el año 2005 y 2010 presentaba una oportunidad relativa de 0,94(IC 95% 0,89-1,00), mostrando un efecto tímidamente protector sobre el fenómeno; pero no fue significativo en el modelo 5.

La variable de segundo nivel que medía el gasto en protección social por unidades de poder de compra por habitante, que fue significativo para hombres y mujeres, solo se asoció positivamente en la categoría que medía entre 6.000,01 y 9.000,00upc/habitante mostrando una proporción protectora del 58% y 53%<sup>15</sup> respectivamente en comparación con los países con un gasto social por debajo de 3.000,00upc/habitante. En el resto de categorías consideradas no se dieron asociaciones significativas.

### 6.2.3. Parámetros aleatorios

El grupo de parámetros aleatorios se ofrecen en la *Tabla 6* y en la *Tabla 7*. Éstos indicaron que un porcentaje de las diferencias entre la población asalariada analizada en estos modelos se podría deberse también a las disimilitudes entre países.

El cálculo del ICC del modelo nulo indicó que un 4,22% de la varianza del estado de salud percibido de los asalariados podría ser explicado por las diferencias entre países, mientras que para mujeres fue del 5,97%.

<sup>15</sup> Este porcentaje protector surge de la ecuación que resta 1 al exponente por ciento:

$$p = (1 - \exp(\beta)) * 100/p = (1 - 0,42) *$$

**Tabla 4. Distribución de la oportunidad relativa en el análisis logístico multinivel para un intervalo de confianza del 95% según el modelo de referencia en Hombres.**

	modelo nulo	modelo 1	modelo 2	modelo 3	modelo 4	modelo 5
Intersección constante	0,24(0,2-0,28)*	0,02 (0,01-0,02)*	0,02 (0,00-0,01)*	0,02 (0,00-0,01)*	0,00 (0,00-0,01)*	0,01 (0,00-0,03)*
Edad continua		1,05 (1,05-1,06)*	1,06 (1,05-1,06)*	1,06 (1,05-1,06)*	1,06 (1,05-1,06)*	1,06 (1,05-1,06)*
Fijo		1	1	1	1	1
Tipo de contrato	Temporal	1,37(1,18-1,57)*	1,25(1,07-1,46)*	1,25(1,07-1,47)*	1,25(1,07-1,46)*	1,25(1,07-1,47)*
	Vía ETI	0,97(0,63-1,47)	0,70(0,44-1,11)	0,69(0,44-1,10)	0,69(0,43-1,10)	0,69(0,44-1,10)
	Aprendizaje o prácticas	4,68(3,15-6,94)*	5,02(3,34-7,56)*	5,10(3,38-7,69)*	5,12(3,39-7,72)*	5,14(3,41-7,76)*
	Sin Contrato	1,21(0,95-1,54)	0,93(0,70-1,24)	0,93(0,70-1,24)	0,94(0,71-1,24)	0,92(0,72-1,27)
Tiempo parcial (Sí)		0,69(0,59-0,81)*	0,83(0,70-0,98)*	0,83(0,70-0,99)*	0,84(0,70-0,99)*	0,84(0,71-0,99)*
Inseguridad laboral (Sí)		1,76(1,46-2,12)*	1,51(1,23-1,84)*	1,52(1,24-1,86)*	1,52(1,25-1,86)*	1,52(1,24-1,86)*
Ex. Psicológicas cuantitativas (Altas)			1,56(1,41-1,73)*	1,56(1,42-1,73)*	1,57(1,42-1,73)*	1,57(1,42-1,74)*
Ex. Psicológicas cognitivas (Altas)			0,96(0,86-1,07)	0,96(0,86-1,06)	0,96(0,86-1,07)	0,96(0,86-1,07)
Previsibilidad (Baja)			1,24(1,11-1,38)*	1,24(1,11-1,37)*	1,24(1,12-1,38)*	1,24(1,12-1,39)*
Control (Alto)			0,80(0,72-0,89)*	0,80(0,72-0,89)*	0,80(0,72-0,89)*	0,80(0,72-0,89)*
Apoyo Social (Alto)			0,87(0,78-0,97)*	0,87(0,78-0,97)*	0,87(0,78-0,97)*	0,87(0,78-0,97)*
Sentimiento del trabajo (Alto)			0,74(0,67-0,83)*	0,74(0,67-0,83)*	0,74(0,67-0,83)*	0,74(0,67-0,83)*

Recompensas (Bajas)				1,94(1,76-2,14)*	1,94(1,76-2,14)*	1,94(1,76-2,14)*	1,94(1,76-2,14)*
Ex. Atención hijos/as (Sí)						1,04 (0,93-1,16)	1,04 (0,93-1,16)
Ex. Mantenimiento hogar y cocinar (Sí)						1,03 (0,92-1,15)	1,03 (0,92-1,16)
Ex. Mayores y/o dependientes (Sí)						0,72(0,51-1,01)	0,71(0,51-1,00)*
Tasa paro 2010							1,01 (0,93-1,10)
Paro larga duración 2010							1,02(0,91-1,13)
Variación tasa paro 2005-2010							0,94 (0,89-1,00)*
Gasto en protección social €año 2010	< 3.000,00						1
	3.000,01-6.000,00						0,84 (0,46-1,52)
	6.000,01-9.000,00						0,42(0,23-0,79)*
	>9.000,01						0,64(0,4-1,19)
Bajo umbral pobreza 2010	< 5,00%						1
	5,01%-10,00%						1,08(0,64-1,80)
	>10,01%						0,70(0,41-1,45)

\*P<0,005. Fuente: Elaboración propia a partir de los datos de la EWCS 2010 facilitados por las ventanas de resultados del procesador SPSS. Se han agrupado en una sola tabla para mejorar su visualización.

**Tabla 5. Distribución de la oportunidad relativa en el análisis logístico multinivel para un intervalo de confianza del 95% según el modelo de referencia en mujeres.**

	modelo nulo	modelo 1	modelo 2	modelo 3	modelo 4	modelo 5
Intersección constante	0,26(0,1-0,32)*	0,03 (0,02-0,04)*	0,01(0,01-0,02)*	0,01(0,01-0,02)*	0,02(0,01-0,04)*	0,03(0,01-0,09)*
Edad continua		1,05 (1,04-1,05)*	1,05 (1,04-1,05)*	1,05 (1,04-1,05)*	1,05 (1,04-1,05)*	1,05 (1,04-1,05)*
Fijo		1	1	1	1	1
Tipo de contrato	Temporal	1,42(1,24-1,63)*	1,30(1,12-1,52)*	1,30(1,11-1,51)*	1,30(1,11-1,51)*	1,30(1,11-1,51)*
	Vía ET	1,34(0,87-2,05)	1,17(0,72-1,88)	1,17(0,72-1,88)	1,17(0,73-1,89)	1,18(0,73-1,90)
	Aprendizaje o prácticas	1,26 (0,77-2,08)	1,49(0,89-2,49)	1,47(0,87-2,47)	1,47(0,87-2,47)	1,48(0,88-2,49)
	Sin Contrato	1,52(1,23-1,87)*	1,33(1,02-1,74)*	1,31(1,01-1,71)*	1,31(1,01-1,71)*	1,33(1,02-1,74)*
	Tiempo parcial (Sí)	0,97(0,88-1,07)	1,03(0,92-1,15)	1,03(0,92-1,16)	1,04(0,93-1,16)	1,04(0,93-1,16)
Inseguridad laboral (Sí)	1,91(1,59-2,29)*	1,59(1,30-1,95)*	1,58(1,29-1,93)*	1,58(1,29-1,93)*	1,58(1,30-1,94)*	1,58(1,29-1,93)*
Ex. Psicológicas cuantitativas (Altas)		1,71(1,54-1,89)*	1,69(1,52-1,88)*	1,69(1,52-1,88)*	1,70(1,53-1,88)*	1,70(1,53-1,89)*
Ex. Psicológicas cognitivas (Altas)		1,06(0,95-1,18)	1,06(0,92-1,18)	1,06(0,92-1,18)	1,06(0,95-1,18)	1,06(0,95-1,18)
Previsibilidad (Baja)		1,17(1,05-1,30)*	1,17(1,05-1,30)*	1,17(1,05-1,30)*	1,17(1,05-1,30)*	1,18(1,05-1,31)*
Control (Alto)		0,78(0,70-0,87)*	0,78(0,70-0,87)*	0,78(0,70-0,87)*	0,78(0,70-0,87)*	0,78(0,70-0,87)*
Apoyo Social (Alto)		0,86(0,77-0,96)*	0,86(0,77-0,96)*	0,86(0,77-0,96)*	0,86(0,77-0,96)*	0,86(0,77-0,96)*
Sentimiento del trabajo (Alto)		0,78(0,69-0,87)*	0,78(0,69-0,87)*	0,77(0,69-0,86)*	0,77(0,69-0,86)*	0,77(0,69-0,86)*

Recompensas (Bajas)				1,94(1,71-2,16)*	1,93(1,73-2,15)*	1,92(1,73-2,14)	1,92(1,73-2,14)
Ex. Atención hijos/as (Sí)					0,89(0,80-1,00)*	0,89(0,80-1,00)*	0,89(0,80-1,00)*
Ex. Mantenimiento hogar y cocinar (Sí)					1,08(0,94-1,24)	1,08(0,94-1,24)	1,08(0,94-1,24)
Ex. Mayores y/o dependientes (Sí)					1,44(1,18-1,77)*	1,44(1,18-1,77)*	1,44(1,17-1,76)*
Tasa paro 2010						0,94(0,82-1,08)	0,94(0,82-1,08)
Paro larga duración 2010						1,04(0,82-1,33)	1,11(0,90-1,36)
Variación tasa paro 2005-2010						0,99(0,92-1,07)	0,98(0,92-1,05)
Gasto en protección social €/año 2010	< 3.000,00						1
	3.000,01						1,02(0,47-2,18)
	-6.000,00						0,47(0,22-0,98)*
	6.000,01-9.000,00						0,54(0,25-1,156)
Bajo umbral pobreza 2010	< 5,00%						1
	5,01%-10,00%						1,13(0,61-2,07)
	>10,01%						0,95(0,39-2,31)

\*P<0,005. Fuente: Elaboración propia a partir de los datos de la EWCS 2010

En los siguientes modelos donde se hizo un análisis más complejo el ICC no se vio muy alterado. El ICC del modelo 4 mostró que las diferencias entre las mujeres fue de 5,92%, porcentaje que bajo al 4,61% en el modelo final al incorporar las variables de protección social y umbral de pobreza. El mismo efecto se observó en los hombres donde la varianza explicada a las diferencias entre países fue de 3,06%, considerando todas las agrupaciones de países<sup>16</sup>.

El parámetro de la MOR comparó la proporción entre un asalariado que pertenecía al país con el mejor estado de salud contra uno del país con el peor estado, pero con las mismas características individuales. Esto vendría a leerse también como lo que variaría el estado de salud si variara el país del individuo.

Este parámetro reflejó que la proporción entre un asalariado del país que tenía el mejor valor del estado de salud contra uno del peor con las mismas características era de 1,44 en hombres y de 1,55 en mujeres en el modelo nulo.

**Tabla 6. Resumen de los parámetros de información según modelo en hombres.**

HOMBRES	Modelo nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Estimación efecto aleatorio(Varianza Nivel 1)	0,145	0,161	0,137	0,139	0,138	0,104
ICC ( <i>thresholdmethod</i> )	4,22	4,67	4,00	4,05	4,03	3,06
Varianza constante $\pi^2/3$	3,29	3,29	3,29	3,29	3,29	3,29
MOR	1,44	1,47	1,42	1,43	1,43	1,36

Fuente: Elaboración propia a partir de los datos de la EWCS 2010

**Tabla 7. Resumen de los parámetros de información según modelos en mujeres.**

MUJERES	Modelo nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Estimación efecto aleatorio(Varianza Nivel1)	0,209	0,219	0,186	0,186	0,207	0,159
ICC ( <i>thresholdmethod</i> )	5,97	6,24	5,35	5,35	5,92	4,61
Varianza constante $\pi^2/3$	3,29	3,29	3,29	3,29	3,29	3,29
MOR	1,55	1,56	1,51	1,51	1,54	1,46

Fuente: Elaboración propia a partir de los datos de la EWCS 2010

<sup>16</sup> A través de este indicador el investigador/a tendría pistas sobre la bondad del modelo y decidir si es más aceptable para su análisis el modelo 4 o 5 y teniendo en cuenta las significaciones de las ORA de los parámetros fijos.

### 6.3. Discusión del caso práctico

Como aportan numerosos estudios de salud laboral las condiciones de empleo se asociaron con el mal estado de salud percibido (Benach et al., 2014). La temporalidad y las consideradas condiciones atípicas de contratación se comportan como un factor de riesgo. Esta situación que expone la precariedad, posiblemente se está extendiendo a todas las formas de ocupación, reforzada en los últimos años por el auge de políticas de alteración de las relaciones laborales y que afectan directamente las condiciones no solo de empleo sino también de trabajo. Sin embargo, cabe destacar que la temporalidad aún tiene una concepción muy diferente entre países: en algunos como Países Bajos es vista como una etapa de inserción, formación y mejora; mientras que en los países del Sur de Europa tiene una connotación asociada a precariedad y vulnerabilidad.

Por otro lado en relación a las condiciones de empleo, la evidencia empírica de multitud de estudios recientes pone de relieve su importancia de las condiciones de trabajo centradas en el entorno psicosocial del trabajo (Kristensen, 2007). Este ámbito está cogiendo mayor interés no solo en la salud laboral y las ciencias de la salud, ya que se está empezando a considerar en otras disciplinas. Existe pero, aún una gran dificultad de homogenización a nivel europeo de los indicadores de medida a pesar de estar reconocida su relación con el estado de salud.

En el caso de las mujeres no se observaron asociaciones al estar mediatizado a nivel europeo y no poderse observar de forma más concreta en este estudio las diferencias entre países. Básicamente podría ser porque los Regímenes de Bienestar han desarrollado de forma desigual la incorporación de las mujeres al mercado de trabajo, las políticas familiares y por ende los modelos de protección global del bienestar. Desde este punto de vista existen diferencias en la población asalariada femenina en relación a las condiciones de empleo y realización del trabajo de la reproducción pero que en este ejercicio no se aprecian.

Por otro lado las variables y covariables de los países se asociaron con el estado de salud. Esta idea ya ha sido suscrita por autoras como Bambra et al. *et al* (2014) también en relación a los efectos de las variables de provisión global del bienestar, y su dimensión explicativa también de las diferencias de salud entre países.

Los parámetros aleatorios complementaron la constatación de las diferencias entre países, que fue el nivel de agrupación de la población asalariada. Aun así, cálculos como el ICC para el modelo final facilitaron unos porcentajes inferiores al 10% que es lo que algunas investigaciones relacionan a nivel europeo. Esto puede ser debido a no utilizar las mismas variables de medida e incluir otras dimensiones en el modelo final.



## 7. CONCLUSIONES

Los modelos multinivel se desarrollaron a partir de variables respuesta de tipo lineal pero los métodos se han generalizado para situaciones donde la respuesta es diferente y no presenta una distribución normal. De esto modo permite trabajar con variables respuesta de tipo binomial y multinomial tan extendidas en las ciencias sociales en general

El modelo logístico multinivel facilita un avance importante en estudios que desean comprobar probabilidades de que un evento suceda o incluso en estudios donde se quiere analizar su contexto.

Los parámetros fijos ayudan a observar las asociaciones de las variables o factores con el fenómeno. A pesar de las dificultades que pueden presentar los parámetros aleatorios, la ICC y la MOR son indicadores bastante gráficos y sustentados por diferentes autores, y ofrecen proporciones para conocer que difiere de un grupo a otro considerando las diferencias. Por tanto un gran complementariedad para los estudios que solo se centran en los parámetros fijos.

Así pues los modelos logísticos conciben la observación de variables de diferentes niveles de agrupación sobre un hecho, y como pueden explicar que parte de la varianza puede ser debido a las diferencias entre sujetos y que parte a los grupos en los que están organizados o anidados.

El análisis de caso con MLR ha sido un ejemplo básico de cómo se puede utilizar esta técnica e ir más allá de la RL tan extendida en ciencias de la salud y donde básicamente se incluyen observaciones individuales. También cómo no, para ver como variables estructurales y de contexto de forma conjunta obtenidas a partir de estadísticas internacionales se pueden incluir en un modelo que dispone de información de una encuesta. Desde este sentido ayuda a inferir la influencia para ayudar a comprender un poco más la variable dependiente.

En este artículo se ha trabajado con un ejemplo dónde se observado diferencias con el mal estado de salud atribuibles a características individuales pero también las disimilitudes de los países en relación al sistema de protección social y el mercado de trabajo. A partir de esta propuesta, el uso aplicado de la MLR podría hacerse extensible a otras preguntas de investigación de las ciencias sociales que pretendan analizar contextos y que trabajen con variables respuesta de tipo categórica.

## 8. LIMITACIONES

Se podría haber realizado el ejercicio en la agrupación del segundo o tercer nivel según Regímenes y así se podrían haber visualizado mejor las disimilitudes de las condiciones de empleo en el caso de las mujeres y su asociación con el mal estado de salud, así como con el resto de covariables. Pero el objetivo de este trabajo no era realizar un análisis de este tipo ni centrarse en los efectos concretos de algunos países a nivel individual, sino mostrar un ejemplo de

modelos multinivel dicotómico con dos niveles dónde el segundo nivel tuviera un número aceptable de clúster.

## 9. BIBLIOGRAFÍA

- ANDRÉU, J. (2011): “EL análisis multinivel: una revisión actualizada en el ámbito sociológico”, *Metodología de Encuestas*, 13, pp. 161–176
- APARICIO, A. y MORERA, M. (2007): La conveniencia del análisis multinivel para la investigación en salud una aplicación para Costa Rica, *Poblacion Y Salud en Mesoamerica*, Vol 4, 2, informe nº 1, disponible en <http://dx.doi.org/10.15517/psm.v4i2.4556> [consulta: 19-07-2015]
- BALLUERKA, N., GOROSTIAGA, A., GÓMEZ-BENITO, J., y HIDALGO, M. D. (2010): “Use of multilevel logistic regression to identify the causes of differential item functioning”, *Psicothema*, 22(4), pp. 1018–1025
- BAMBRA, C., y EIKEMO, T. A. (2009): “Welfare state regimes, unemployment and health: a comparative study of the relationship between unemployment and self-reported health in 23 European countries”, *Journal of Epidemiology and Community Health*, 63(2), pp. 92–98
- BAMBRA, C., LUNAU, T., VAN DER WEL, K. A., EIKEMO, T. A., y DRAGANO, N. (2014): “Work, health, and Welfare: The association between Working Conditions, Welfare States, and self-reported Health in Europe”, *International Journal of Health Services*, 44, pp. 113–136
- BENACH, J. y MUNTANER, C. (2010): Empleo, trabajo y desigualdades en salud: una visión global, Barcelona, Icaria
- BENACH, J., TARAFÁ, G. y RECIO, A. (2014): Sin trabajo, sin derechos, sin miedo, Barcelona, Icaria
- BERKHOF, J. y SNIJDERS, T. A.B. (2001): “Variance Component Testing in Multilevel Models”, *Journal of Educational and Behavioral Statistics*, 24 (2) pp. 133–152
- BLIESE, P.D. y JEX, S. M. (2002): “Incorporating a multilevel perspective into occupational stress research: Theoretical, methodological, and practical implications”. *Journal of Occupational Health Psychology*, 7, pp. 265–276.
- CAMPOS-SERNA, J., RONDA-PÉREZ, E., MOEN, B. E., ARTAZCOZ, L., y BENAVIDES, F. G. (2013): “Welfare state regimes and gender inequalities in the exposure to work-related psychosocial hazards”, *International Journal of Occupational and Environmental Health*, 19(3), PP. 179–95
- CATALÁN-REYES, M., y GALINDO-VILLARDÓN, M. (2003): “Utilización de los modelos multinivel en investigación sanitaria”, *Gaceta Sanitaria*, 17, 3, pp. 35–52
- CLARK, T. y LINZER, D.A. (2015): “Should I use Fixed or Random Effects?”, *Political Science Research and Methods*, 3, 2, pp. 399–408
- CEBOLLA BOADO, H. (2013): Introducción al análisis multinivel, Madrid, Centro de Investigaciones Sociológicas (CIS)
- COURGEAU, D. (Ed.) (2008): Multilevel modeling of educational data. Methodology and Epistemology of Multilevel Analysis, Springer Netherlands, Rotterdam
- EIKEMO, T. A., BAMBRA, C., JUDGE, K., y RINGDAL, K. (2008): “Welfare state regimes and differences in self-perceived health in Europe: A multilevel analysis”, *Social Science and Medicine*, 66(11), PP. 2281–2295

- HECK, R. H., y THOMAS, S. L. (2009): *An introduction to multilevel modeling techniques*, New York, Routledge
- HECK, R. H., THOMAS, S. L., y TABATA, L. N. (2012): *Multilevel modeling of categorical outcomes using IBM SPSS*, New York, Routledge
- HOX, J.J. (2002): *Multilevel analysis: Techniques and applications*, New York, Lawrence Erlbaum Associates
- HUIJTS, T., PERKINS, J. M., y SUBRAMANIAN, S. V. (2010): Political regimes, political ideology, and self-rated health in Europe: A multilevel analysis. *PLoS ONE*, 5(7)., disponible en <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2908625/pdf/pone.0011711.pdf> [consulta: 19-07-2015]
- JOVELL, A.J. (2006): *Análisis de regresión logística*, Madrid, Centro de Investigaciones Sociológicas.
- KARASEK, R., y THEORELL, T. (1999): *Healthy work. Stress, productivity, and reconstruction of working life*, New York, N. B. Books, Ed
- KRISTENSEN, T. S. (2007): "El ambiente psicosocial del trabajo y la salud", En *La población asalariada y la salud*, Barcelona, ISTAS, pp. 9–32
- MANOR, O., MATTHEWS, S. y POER, C. (2000): "Dichotomous or categorical response? Analysing self-rated health and lifetime social class", *International journal of epidemiology*, 29 (1), pp.149-157
- MERLO, J., et al. (2006): "A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena", *Journal of Epidemiology and Community Health*, 60(4), pp.290–297
- MONCADA, S., UTZET, M., MOLINERO, E., LLORENS, C., MORENO, N., GALTÉS, A., y NAVARRO, A. (2014): "The copenhagen psychosocial questionnaire II (COPSOQ II) in Spain-A tool for psychosocial risk assessment at the workplace", *American Journal of Industrial Medicine*, 57, pp. 97–107
- MUNTANER, C. et al. (2011): "Politics, welfare regimes, and population health: controversies and evidence" *Sociology of Health & Illness*, 33 (6), pp.946-964
- MURILLO TORRECILLA, F. J. (2008): "Los modelos multinivel como herramienta para la investigación educativa", *Magis, Revista Internacional de Investigación Educativa*, 1, pp.45–62.
- PARDO, A., ÁNGEL, M., SAN, R. (2007): "Cómo ajustar e interpretar modelos multinivel con SPSS", *Psicotherma*, 19(2), pp. 308–321
- PÉREZ FERNÁNDEZ, V. (2012): *Los modelos multinivel en el análisis de factores de riesgo de síndromes recurrentes en lactantes. Enfoques frecuentista y bayesiano*, Universidad de Murcia, disponible en <http://tesisenred.net/handle/10803/109213>, [consulta: 19-09-2014]
- RABE-HESKETH, S., y SKRONDAL, A. (2006): "Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society, Journal of the Royal Statistics in Society: Series A*, vol 169, 4, pp. 805-827
- SALANOVA SORIA, M., y SCHAUFELI, W. B. (2002): "¿Cómo evaluar los riesgos psicosociales en el trabajo? Prevención", *Trabajo Y Salud: Revista Del Instituto Nacional de Seguridad E Higiene En El Trabajo*, 20 , pp. 4–10.
- SIEGRIST, J. (2011): "Social determinants of health: contributions from European health and medical sociology", *Política Y Sociedad*, 48(2), pp. 249–258
- SNIJDERS, T.A.B. y BOSKER, R. J. (1999): *Analysis: An Introduction to Basic and Advanced Multilevel Modelling*, London, Sage Publishers
- VAN YPEREN, N. W., y SNIJDERS, T.A. B. (2000): "A Multilevel Analysis of the Demands--Control Model: Is Stress at Work Determined by Factors at the Group

Level or the Individual Level?“, Journal of Occupational Health Psychology, 5 (1) PP. 182-190

## 10. ANEXOS

**Tabla 8. Distribución de la población asalariada incluida en el análisis según país y sexo.**

País	Hombres	Mujeres	País	Hombres	Mujeres
Bélgica	351	304	Luxemburgo	19	14
Bulgaria	260	240	Hungría	294	277
República checa	367	320	Malta	16	9
Dinamarca	225	220	PP. Bajos	660	602
Alemania	3233	2801	Austria	319	292
Estonia	42	50	Polonia	1067	1010
Grecia	296	224	Portugal	379	335
España	1611	1281	Romania	625	507
Francia	1932	1903	Eslovenia	71	67
Irlanda	142	136	Eslovaquia	193	159
Italia	1830	1300	Finlandia	186	191
Chipre	29	26	Suecia	355	347
Letonia	76	85	Reino Unido	2204	2138
Lituania	106	110	Total	16891	14948

Fuente: Elaboración propia a partir de los datos de la EWCS 2010

**Tabla 9. Análisis descriptivo de las variables incorporadas en el estudio de las frecuencias.**

	Hombres		Mujeres	
	categoría base %	categoría de contraste %	categoría base %	categoría de contraste %
Estado de salud percibido	Muy buena y buena (80,2)	Pasable-regular, mala y muy mala (19,8)	Muy buena y buena (78,2)	Pasable-regular, mala y muy mala (21,8) (19,8)
Edad	Edad media (40,1, des. St. 11,8)		Edad media (39,93, des. St. 11,5)	
Tipo de contrato	Fijo (81)	Temporal (10,7) Vía ETT (1,4) Aprendizaje (1,0) sin contrato (4,3)	Fijo (77,3)	Temporal (13,4) Vía ETT (1,2) Aprendizaje (1,2) Sin contrato (5,1)
Tiempo parcial	Sí (10,8)	No (89,2)	Sí (38,3)	No (61,7)
Inseguridad laboral	Muy de acuerdo (5,1)	Resto categorías (94,9)	Muy de acuerdo (5,6)	Resto categorías (94,4)
Ex. Psicológicas cuantitativas	Altas (57,1)	Bajas (42,9)	Altas (64,4)	Bajas (35,6)
Ex. Psicológicas cognitivas	Altas (66,7)	Bajas (33,3)	Altas (60,84)	Bajas (39,16)
Previsibilidad	Alta (29,1)	Baja (70,9)	Alta (33,96)	Baja (66,04)
Control	Alto (45,9)	Bajo (54,1)	Alto (47,98)	Bajo (52,02)
Apoyo Social	Alto (32,85)	Bajo (77,15)	Alto (34,97)	Bajo (65,03)
Sentimiento del trabajo	Alto (56,79)	Bajo (43,21)	Alto (41,40)	Bajo (58,
Recompensas	Altas (47,86)	Bajas (52,14)	Altas (53,64)	Bajas (46,46)
Ex. Atención hijos/as	Sí (25,4)	No (74,8)	Sí (40,8)	No (59,2)
Ex. Mantenimiento hogar y cocinar	Sí (25,0)	No (75,0)	Sí (77,1)	No (32,9)
Ex. Atención gente mayor i/o dependientes	Sí (2,1)	No (97,9)	Sí (5,2)	No (94,8)

Fuente: Elaboración propia a partir de los datos de la EWCS 2010

**Tabla 10. Comparación de la ORA del modelo final según la Regresión Logística Multinivel (MLR) y la Regresión Logística Multivariada (LR).**

	Hombres		Mujeres	
	MLR	LR	MLR	LR
Intersección LRM / constante RLM (binaria)	0,01 (0,00-0,03)*	0,19*	0,03(0,01-0,09)*	0,16*
Edad continua	1,06 (1,05-1,06)*	1,06(1,06-1,07)*	1,05 (1,04-1,05)*	1,05(1,04-1,05)*
Tipo de contrato	Fijo	1	1	1
	Temporal	1,25(1,07-1,47)*	1,24(1,06-1,45)*	1,32(1,13-1,53)*
	Vía ETT	0,69(0,44-1,10)	0,68(0,43-1,09)	1,12(0,7-1,8)
	Aprendizaje o prácticas	5,14(3,41-7,76)*	4,84(3,2-7,33)*	1,26(0,72-2,2)
	Sin Contrato	0,92(0,72-1,27)	0,9(0,68-1,18)	1,17(0,89-1,53)
Tiempo parcial (Sí)	0,84(0,71-0,99)*	0,88(0,74-1,05)*	1,04(0,93-1,16)	0,99(0,88-1,1)
Inseguridad laboral (Sí)	1,52(1,24-1,86)*	1,52(1,25-1,85)*	1,58(1,29-1,93)*	1,48(1,21-1,81)*
Ex. Psicológicas cuantitativas (Altas)	1,57(1,42-1,74)*	1,51(1,36-1,67)*	1,70(1,53-1,89)*	1,61(1,45-1,79)*
Ex. Psicológicas cognitivas (Altas)	0,96(0,86-1,07)	0,97(0,87-1,08)	1,06(0,95-1,18)	1,02(0,91-1,14)
Previsibilidad (Baja)	1,24(1,12-1,39)*	1,21(1,08-1,34)*	1,18(1,05-1,31)*	1,21(1,09-1,35)*
Control (Alto)	0,80(0,72-0,89)*	0,82(0,74-0,91)*	0,78(0,70-0,87)*	0,8(0,72-0,89)*
Apoyo Social (Alto)	0,87(0,78-0,97)*	0,85(0,76-0,95)*	0,86(0,77-0,96)*	0,83(0,74-0,93)*
Sentimiento del trabajo (Alto)	0,74(0,67-0,83)*	0,77(0,69-0,86)*	0,77(0,69-0,86)*	0,8(0,71-0,89)*
Recompensas (Bajas)	1,94(1,76-2,14)*	2,03(1,84-2,24)*	1,92(1,73-2,14)*	1,96(1,76-2,18)*
Ex. Atención hijos/as (Sí)	1,04 (0,93-1,16)	1,06(0,95-1,19)	0,89(0,80-1,00)*	0,94(0,84-1,05)
Ex. Mantenimiento hogar y cocinar (Sí)	1,03 (0,92-1,16)	1,01(0,9-1,13)	1,08(0,94-1,24)	1,09(0,96-1,25)

Ex. Mayores y/o dependientes (Sí)		0,71(0,51-1,00)*	0,81(0,59-1,13)	1,44(1,17-1,76)*	1,46(1,19-1,79)*
Tasa paro 2010		1,01 (0,93-1,10)	1(0,97-1,04)	0,94(0,82-1,08)	1,01(0,95-1,07)
Paro larga duración 2010		1,02(0,91-1,13)	1,09(1,05-1,14)*	1,11(0,90-1,36)	1,01(0,93-1,09)
Variación tasa paro 2005-2010		0,97(0,92-1,03)	0,97(0,96-0,99)*	0,98(0,92-1,05)	0,98(0,96-1)
Gasto en protección social €/año 2010	< 3.000,00	1	1	1	1
	3.000,01 -6.000,00	0,84 (0,46-1,52)	0,64(0,52-0,8)*	1,02(0,47-2,18)	0,55(0,43-0,7)*
	6.000,01-9.000,00	0,42(0,23-0,79)*	0,4(0,32-0,5)*	0,47(0,22-0,98)*	0,29(0,23-0,37)*
Bajo umbral pobreza 2010	>9.000,01	0,64(0,4-1,19)	0,48(0,39-0,61)*	0,54(0,25-1,156)	0,34(0,26-0,44)*
	< 5,00%	1	1	1	1
	5,01%-10,00%	1,08(0,64-1,80)	0,93(0,76-1,13)	1,13(0,61-2,07)	1,12(0,91-1,37)
	>10,01%	0,70(0,41-1,45)	0,59(0,46-0,75)*	0,95(0,39-2,31)	0,64(0,48-0,87)*

\*P<0,005. Fuente: Elaboración propia a partir de los datos de la EWCS 2010 facilitados por las ventanas de resultados del procesador SPSS. Se han agrupado en una sola tabla para mejorar su visualización.



