Mathematical Model for Estimating Nutritional Status of the Population with Poor Data Quality in Developing Countries: The Case of Chile

Denisse Ávalos[1], Cristóbal Cuadrado[2,3][a], Jocelyn Dunstan[4,5][b], Javier Moraga-Correa[1,6], Luis Rojo-González[1,7][c], Nelson Troncoso[1] and Óscar C. Vásquez[1][d]

[1]*Department of Industrial Engineering, Universidad de Santiago, Santiago, Chile*
[2]*School of Public Health, Universidad de Chile, Santiago, Chile*
[3]*Centre for Health Economics, University of York, York, U.K.*
[4]*Center for Mathematical Modeling - CNRS UMI2807, Universidad de Chile, Santiago, Chile*
[5]*Center for Medical Informatics and Telemedicine, Universidad de Chile, Santiago, Chile*
[6]*Business School, University of Nottingham, Nottingham, U.K.*
[7]*Facultat de Matemàtiques i Estadística, Universitat Politècnica de Catalunya, Barcelona, Spain*

Keywords: Transition Probabilities, Obesity, Developing Countries, Non-linear Programming, Poor Data Quality.

Abstract: Obesity is one of the most important risk factors for non-communicable diseases. Nutritional status is generally measured by the body mass index (BMI) and its estimation is especially relevant to analyse long-term trends of overweight and obesity at the population level. Nevertheless, in most context nationally representative data on BMI is scarce and the probability of individuals to progress to obese status is not observed longitudinally. In the literature, several authors have addressed the problem to obtain this estimation using mathematical/computational models under a scenario where the parameters and *transition probabilities* between nutritional states are possible to compute from regular official data. In contrast, the developing countries exhibit poor data quality and then, the approaches provided from the literature could not be extended to them. In this paper, we deal with the problem of estimating nutritional status transition probabilities in settings with scarce data such as most developing countries, formulating a non-linear programming (NLP) model for a disaggregated characterization of population assuming the transition probabilities depend on sex and age. In particular, we study the case of Chile, one of the countries with the highest prevalence of malnutrition in Latin America, using three available National Health Surveys between the years 2003 and 2017. The obtained results show a total absolute error equal to 5.11% and 10.27% for sex male and female, respectively. Finally, other model applications and extensions are discussed and future works are proposed.

1 INTRODUCTION

The nutritional status of the population is generally measured using the body mass index (BMI). This metric is defined as the weight divided by height squared, kg/m^2 (Okorodudu et al., 2010; Apovian, 2016). In contrast to other technical metrics such as dual-energy x-ray absorptiometry, BMI is easy-to-implement in clinical practice and population surveys and provides a similar discriminatory capability (Huxley et al., 2010; NCD Risk Factor Collaboration (NCD-RisC), 2016; NCD Risk Factor Collaboration (NCD-RisC), 2017). In practice, this metric allows to classify the nutritional status for individuals aged 18 years and older according to its BMI into four groups proposed by the World Heath Organization (WHO): i) Underweight, BMI < 18.4; ii) Normal, $18.5 \leq$ BMI ≤ 24.9; iii) Overweight, $25.0 \leq$ BMI ≤ 29.9; and iv) Obese, BMI ≥ 30.

Unfortunately, obesity represents one of the most important risk factors for non-communicable diseases. It was declared an epidemic in 1990 by the World Health Organization, reaching the first position in the 21st century. Between the years 1995 and 2000,

[a] https://orcid.org/0000-0002-0174-5958
[b] https://orcid.org/0000-0001-6726-7242
[c] https://orcid.org/0000-0003-4894-2470
[d] https://orcid.org/0000-0002-1393-4692

Ávalos, D., Cuadrado, C., Dunstan, J., Moraga-Correa, J., Rojo-González, L., Troncoso, N. and Vásquez, Ó.
Mathematical Model for Estimating Nutritional Status of the Population with Poor Data Quality in Developing Countries: The Case of Chile.
DOI: 10.5220/0010262404080415
In *Proceedings of the 10th International Conference on Operations Research and Enterprise Systems (ICORES 2021)*, pages 408-415
ISBN: 978-989-758-485-5

adult population with obesity increased in 100 million people, reaching 300 million worldwide, whereas 18 million under-five year children have overweight (World Health Organization, 2003). In addition, more than 1.9 billion adults aged 18 years and older were overweight, being 650 million adults were obese in 2016 (World Health Organization, 2016).

Consequently, the estimation of the prevalence of different nutritional status categories is especially relevant for the development of future health policies due to at least two reasons:

- Obesity is associated with the leading causes of death worldwide since the management of its related factors has proven to be tremendously complex presenting a sustained upward trend. In particular, obesity is related to serious health risks (National Institutes of Health, 1998). For instance, every 5-unit increases in BMI above 25, the overall mortality increases by 29%, cardiovascular mortality by 41%, and diabetes-related mortality by 210% (Apovian, 2016).
- In terms of the associated costs, a person who has obesity incurs an annual healthcare cost of 36% greater than a normal weight person and a 77% higher in the medication costs (Sturm, 2002). In this sense, some authors have been reported that an obese patient has annual medical spending between 25% (Detournay et al., 2000) and 42% (Finkelstein et al., 2003) higher than who is of normal weight. This account for direct costs of, at least, 5%, of the total health expenditures of developed countries (Levy et al., 1995; Manson et al., 2004). In practice, if the number of individuals aged 16 and 17 who have overweight or obesity could be reduced by 1%, this would result in a decrease in lifetime medical costs of $586 million (Wang et al., 2010). In addition, it might be expected that at low-income levels of the country the underweight prevalence should dominate the landscape, but the projected average annual growth of obesity indicates that, across 147 countries, it will increase 2.47% respect to the Gross Domestic Product per capita during 2019-2024 (Talukdar et al., 2020).

The problem of estimation of long-term trends of obesity prevalence is mainly addressed by using mathematical models that aim to capture the relationships among environment, biological, social and cultural aspect from a system point of view under a scenario where regular official data is available (Mitchell et al., 2011; Frerichs et al., 2013). However, these approaches could not be easily extended to other settings where the data is not collected regularly or not extensively enough. For instance, the developing countries such as Argentina (Secretaría de Gobierno de Salud, Argentina, 2019), Brazil (Instituto Brasileiro de Geografía e Estadística, Brasil, 2020), Colombia (Ministerio de Salud, Colombia, 2015) and Chile (Ministerio de Salud, Chile, 2017) only provide two and three national health survey into two decades.

1.1 Literature Review

In the literature, the importance of estimating obesity prevalence over time and how to use National Health Surveys in each country as main data source, is emphasized from different modeling perspectives (see (Olariu et al., 2017) for a recent survey in Markov cohort models and (Xue et al., 2018) for a recent survey in system dynamic (SDM) and agent-based modeling (ABM)). In particular, the researchers mention a special key input called *transition probabilities* in the chronic disease modeling context. They are defined as the probability or rate to move from one state of a categorical risk factor (e.g. nutritional status) to another (Van de Kassteele et al., 2012); being *transition rates* and *transference rates* sometimes used as synonyms.

In the last decades, several authors have addressed the estimation of the nutritional status for the population from developed countries. For instance, a multi-stage Markov model is decribed in (Van de Kassteele et al., 2012), which is used on the "Permanent Onderzoek Leef Situatie" data for the Netherlands collected between 2006 and 2007, that is to say, a cross-sectional study, considering the individuals of 85 years old tops. To compute the transition probabilities the authors consider a transportation problem (well-known in Linear Programming), where the results show that the prevalence of being normal weight during the last year and have obesity the next year, and vice versa, is null at any age. A susceptible–infected–recovered model for the United States (US) and the United Kingdom (UK) is proposed in (Ward et al., 2017), using the information between 1988 and 1998 (Ogden et al., 2006), and the Health Survey for England between 1993 and 1997 (Thomas et al., 2014), respectively; indicating that, via a forward simulation, the obesity prevalence will plateau independent of current prevention strategies at 32% (US) and 39.6% (UK) by about 2030. A longitudinal study, from an ABM simulation point of view, was carried out using the National Health and Nutrition Examination Survey from 1976 through 2014 to predict risk factors for the nutritional statuses for individuals aged 35 years. The results project that a majority of children (57.3%) will be obese at the age of 35 years, and roughly half of the projected prevalence will occur during childhood.

Recently, several authors have tackled the problem of estimating the nutritional status of the population with poor data quality. A time-homogeneous continuous-time Markov model is proposed in (Lartey et al., 2020) to compute the transition probabilities between nutritional statuses, using the Ghana Who SAGE in 2007/2008 and 2014/2015. The obtained results show that, for obese individuals, the probability of remaining obese, decrease to overweight and normal weight was 90.2%, 9.2%, and 0.6%, respectively.

Complementary to the previous works in this field, our research acknowledges the reality of the so-called developing countries, where the data is not collected regularly or not very extensively.

1.2 Our Contribution

This research deals with the problem of estimating the nutritional status of the population with poor data quality in developing countries. In particular, we study the case of Chile, which ranks third in Latin America in child overweight prevalence with a presence of 9.3% (Food and Agriculture Organization, 2018) and second among the countries of the Organization for Economic Cooperation and Development in overweight prevalence for the population over 15 years old. Chile represents an example where the data is not collected regularly or not very extensively, as it is often the case in other regions of the 'developing' world. Formally, we formulate a non-linear programming (NLP) model, which allows us to determine the *transition probabilities* considering a set of disaggregated variables by BMI and age ranges and assuming the transition probabilities depend on sex and age range. The parameters are obtained from the analysis of the official reports and the three available national health surveys between the years 2003 and 2017. To test the performance of our NLP model, we carry out computational experiments and compute the total absolute error of our estimation in different scenarios. Finally, other applications for the model are discussed and future works are proposed.

2 NON-LINEAR PROGRAMMING (NLP) MODEL

2.1 Dynamic Conceptualization

In order to define our model, we introduce the dynamic conceptualization for the population's nutritional status during a given period. Let $\mathcal{I}$, $\mathcal{J}$ and $\mathcal{K}$ be the sets of discretized BMI, age ranges, and sexes, respectively. For convenience, we denote $\mathcal{T}$ the set of the years within the given period and the state of the population is defined by BMI i, age range j, and sex k. From each state, the movements within and between the BMI and age ranges are modeled according to the transition probabilities. We distinguish the transition probabilities to decrease, to remain or to increase the current BMI as follow: i) $\alpha_{j,k}$, is the transition probability to increase the BMI of the people of sex k from the year t to the year $(t+1)$; ii) $\beta_{j,k}$, is the transition probability to decrease the BMI of the population of sex k from one year t to the year $(t+1)$ and iii) $\phi_{j,k}$, is the transition probability to remain the BMI of the population of sex k from one year t to the year $(t+1)$.

In addition, we consider the population growth $\kappa^t_{i,j,k}$ and the population proportion $\gamma^t_{j,k}$ of sex k in the upper bound of the age range j at year t (i.e., the population proportion that would change their age range from year t to year $(t+1)$). Note that this proportion is assumed equal to 0% in the last age range. Thus, the transition probabilities for the above population proportion can be separated into two sub-categories as follow: i) $\eta^{\alpha,t}_{j,k} := \gamma^t_{j,k}\alpha_{j,k}$, $\eta^{\beta,t}_{j,k} := \gamma^t_{j,k}\beta_{j,k}$ and $\eta^{\phi,t}_{j,k} := \gamma^t_{j,k}\phi_{j,k}$ as the transition probabilities to increase, to decrease or to remain the BMI from year t to the year $(t+1)$ but, at the same time, changing their age range; and ii) $\psi^{\alpha,t}_{j,k} := (1-\gamma^t_{j,k})\alpha_{j,k}$, $\psi^{\beta,t}_{j,k} := (1-\gamma^t_{j,k})\beta_{j,k}$ and $\psi^{\phi,t}_{j,k} := (1-\gamma^t_{j,k})\phi_{j,k}$ as the transition probabilities to increase, to decrease or to remain the BMI from year t to the next one $(t+1)$ but, at the same time, remaining their age range. An illustration of the dynamic conceptualization is shown in Figure 1.

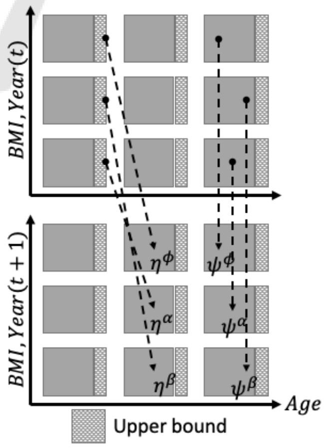

Figure 1: Illustration of the dynamic conceptualization. Two consecutive years, t and $t+1$ are depicted in order to show the transitions of some defined population.

2.2 Formulation

We denote $I^t_{i,j,k}$ the continuous positive variable that represents the population of BMI $i \in I$, age range $j \in \mathcal{J}$ and sex $k \in \mathcal{K}$ at year $t \in \mathcal{T}$. Let $\underline{T}$ and $\overline{T}$ be the bounds on the years in the model. The model aims to minimize the weighted mean squared error in the fitting of the population during the period in $[\underline{T};\overline{T}]$. We define $w^{\overline{T}}_{j,k}$ as the population proportion of sex $k \in \mathcal{K}$ in age range $j \in \mathcal{J}$ at the end of the year $\overline{T}$. Let $U^{\overline{T}}_{i,j,k}$ be the real population at the end year $\overline{T}$. Given a sex $k \in \mathcal{K}$, the NLP model is defined as follows:

$$[\text{MIN}] \sum_{i \in I} \sum_{j \in \mathcal{J}} (I^{\overline{T}}_{i,j,k} - U^{\overline{T}}_{i,j,k})^2 w^{\overline{T}}_{j,k} \tag{1}$$

subject to (for each year $t \in \{\underline{T} \dots \overline{T} - 1\}$)

$$\begin{aligned}
\frac{I^{t+1}_{i,j,k}}{1-\kappa^t_{i,j,k}} =& \psi^{\alpha,t}_{j,k} I^t_{i-1,j,k} + + \eta^{\alpha,t}_{j-1,k} I^t_{i-1,j-1,k} + \psi^{\phi,t}_{j,k} I^t_{i,j,k} \\
& + \eta^{\phi,t}_{j-1,k} I^t_{i,j-1,k} + \psi^{\beta,t}_{j,k} I^t_{i+1,j,k} + \eta^{\beta,t}_{j-1,k} I^t_{i+1,j-1,k} \\
& \forall j \in \mathcal{J} \setminus \{1\}, \forall i \in I \setminus \{1, |I|\}
\end{aligned} \tag{2}$$

$$\gamma^t_{j,k} = \eta^{\alpha,t}_{j,k} + \eta^{\beta,t}_{j,k} + \eta^{\phi,t}_{j,k} \qquad \forall j \in \mathcal{J} \tag{3}$$

$$(1 - \gamma^t_{j,k}) = \psi^{\alpha,t}_{j,k} + \psi^{\beta,t}_{j,k} + \psi^{\phi,t}_{j,k} \qquad \forall j \in \mathcal{J} \tag{4}$$

$$\alpha^{min}_{j,k} \leq \eta^{\alpha,t}_{j,k} + \psi^{\alpha,t}_{j,k} \leq \alpha^{max}_{j,k} \qquad \forall j \in \mathcal{J} \tag{5}$$

$$\beta^{min}_{j,k} \leq \eta^{\beta,t}_{j,k} + \psi^{\beta,t}_{j,k} \leq \beta^{max}_{j,k} \qquad \forall j \in \mathcal{J} \tag{6}$$

$$\phi^{min}_{j,k} \leq \eta^{\psi,t}_{j,k} + \psi^{\psi,t}_{j,k} \leq \phi^{max}_{j,k} \qquad \forall j \in \mathcal{J} \tag{7}$$

$$I^t_{i,j,k}, \eta^{\alpha,t}_{j,k}, \eta^{\beta,t}_{j,k}, \eta^{\phi,t}_{j,k}, \psi^{\alpha,t}_{j,k}, \psi^{\beta,t}_{j,k}, \psi^{\phi,t}_{j,k} \geq 0 \qquad \forall i \in I, j \in \mathcal{J} \tag{8}$$

Expression (1) states the objective function of the model. The possible transition is defined by the set of constraints (2). These transitions are constrained to the proportion to change of age range given by the set of constraints (3)–(4), the bounds on these transitions are imposed by the set of constraints (5)–(7) and the domain of the variables is defined in the set of constraints (8).

2.3 Discretization and Data Wrangling

To compute the required parameters, we consider official reports such as the three collected National Health Surveys (ENS) into 15 years provided by the Ministry of Health of Chile (MINSAL) (Ministerio de Salud, Chile, 2017), the governmental documents from the Department of Statistics and Health Information (DEIS) (Departamento de Estadística e Información de Salud, 2018) and the National Institute of Statistics (INE) (Instituto Nacional de Estadística, 2017). Since the proposal required a discretization of the variables, we should do it to achieve the desired shape of the input. Thus, we consider the BMI as the integer value of the recorded value and to group the age by ranges of 10 years starting from 15 years old such as a population grid is obtained.

However, as the survey is only a sample of the entire population, there would likely be some combinations of the grid that have zero-counting. Therefore, we should avoid this first issue by redistributing the population. So, the discussion now turns into which model to pick and how to set its parameters. Even although we should deal with a bi-variate distribution over the BMI and the age range, there is not evidence of this, but it does by considering the BMI itself. For instance, it is common to assume a Gaussian distribution, but it would be true if and only if the related multi-factorial processes have additive effects; nonetheless, for biologic variables, the knowledge suggests that these processes have a multiplicative effect, e.g., an obesogenic environment, which is more likely to follow a skewed, possibly log-normal, distribution (Penman and Johnson, 2006).

A particular analysis is carried out on the population growth, $\kappa^t_{i,j,k}$. The population growth is the difference between the birth rate and the mortality rate and, as we are considering people older than 15 years old, it is necessary to get this rate in a time-lagged sense, i.e., we need the birth rate of the people that were born 15 years ago. Besides, a requirement is that this rate should be distributed by BMI and age ranges. Addressing this latter issue, from a separable point of view, we have that: i) the lagged birth rate only influences on the first age range whereas the mortality rate is in the whole set of age ranges; ii) to get the distributed population growth, we use a study related to survival analysis which works on the *Hazard ratios* (Berrington de González et al., 2010) to get them by BMI.

Even although the transition probabilities are the core of the study itself, we must provide the feasible bounds for the estimation. Several authors have been reported evidence on the transition probabilities for a specific nutritional status (Power et al., 1997; Orpana et al., 2006; Laitinen et al., 2001; Booth et al., 2012; Fildes et al., 2015; Srinivasan et al., 1996). However, the obtained results reported by them are not necessarily the BMI and age range considered. To address this problem, we propose to work with multiple scenarios approach from the literature data. Thus, the bounds of the transition probabilities, X, and its mean value, $Ave(X)$, are assumed to be normal distributed such that the confidence bounds of the transition probability, say $\underline{Ave(X)}$ as the lower bound, and $\overline{Ave(X)}$ as the upper bound, at a confidence level $\delta\%$ can be obtained from $Pr(\underline{Ave(X)} \leq X \leq \overline{Ave(X)}) = \delta$.

3 RESULTS

The computational experiments are carried out splitting up the data into a fitting period (2003-2010) and a validation period (2010-2017) as a forward simulation to set out the performance of the model.

3.1 Implementation

The mathematical model is implemented in AMPL programming language with Minos as the solver and executed on a MacBook Air Intel i5, 1.6 GHz, 8GB RAM, with no stopping time, 2500 iterations tops considering five random seeds given the non-linearity of the model. The obtained results are such that we use to make a discrete grid at confidence levels in $[0.01; 0.995]$ such as 256 and 254 scenarios are given with solving time of 38.96 minutes and 41.90 minutes for males and females, respectively.

3.2 Transition Probabilities

The obtained results show that there are some differences between the sexes. In particular, such as Table 1 shows, we can see that the standard deviation for each fitted parameter is quite similar between sexes and the comparison between them shows a remarkable difference over the age ranges but not at all. For instance, while the transition probability to remain the current BMI is the highest one, the transition probability to increase the BMI is greater than to decrease until the 65 years old for both sexes but between 65 and 74 years old, the sex female reverses this trend, i.e., the transition probability to increase the BMI is greater than to decrease, then this relation gets reverse and then goes back to the same.

A second important remark concerns the density of the results (see Figure 2). In fact, the transition probabilities are ever in the same order; in decreasing order, they are to remain, to increase, and to decrease. So, in general, although the transition probability to remain the current BMI dominates the others, this effect does not make the difference due to remain a status is possible at any level and if we compare from one year to another it means that the population has a stable BMI, but when we make the period wider, e.g., 2 years, if the person increases its BMI in the first place and then remains it, the overall effect is that the person increases its BMI from the beginning of the study. Therefore, it is clear that the population, in the long-term sense, has a trend to increase their BMI, but important differences especially exist for sex male and some age ranges, including ones where the transition probabilities to decrease or to increase are almost the same.

For both sexes there is an age range that changes with respect to the others, for sex male, it is 55-64 years old and for sex female, it is 15-24 years old. These age ranges have completely different shapes of the distribution; in fact, they are, from a biological point of view, interesting phases of hormonal changes in the body, so they should be carefully analyzed. On the other hand, for both sexes, a useful remark is related to the variability and the shape of the transition probabilities. The distributions are wide enough to cover the half of the probabilities, at least, whereas there are not bi-modal ones except the distributions for the first age range. However, the entire previous comments are just true if the fitting is well. Using the total absolute error as a metric to compare the forecasting population against the real population in 2017, it reaches a 5.11% and 10.27% for sex male and female, respectively. Even when the numbers seem quite well, it is necessary to see the behavior of the estimation. Figure 3 depicts, by sex, the real population at each BMI (points) and the fitted curve obtained from the model (lines) including their respectively 95% confidence intervals at the end of the fitting period (2010) and at the end of the validation period (2017). The comparison shows quite similar behavior and, with low total absolute error for both sexes, and the fact that the real data belongs to the confidence intervals (which are not too wide), the fitting might be understood as a good one but taking into account that both periods differ from each other.

4 DISCUSSION AND FUTURE WORK

This work addressed the problem of estimating nutritional status with poor data quality. A non-linear programming model is formulated to obtain the transition probabilities that are key elements in the modeling of this epidemic and the designing of malnutrition public policy options and the computation of the associated costs of a specific population. Our proposal provides a novel methodology to estimate them with a few data available and a disaggregated characterization of the population by sex, BMI and age ranges, in contrast to other similar work works.

In particular, the research was focused on the case of Chile, one of the countries with the highest malnutrition in Latin America. The obtained results show the fitted transition probabilities are fair via a straight comparison against the last known information. Nonetheless, it is important to say that, despite

Table 1: Summary of the transition probabilities by sex and age ranges considering the whole fitting period (2003-2010). The standard deviation (*Sd*) for each fitted value is similar between sexes. The comparison, in mean ($\bar{x}$) terms, shows that the transition probability to increase the BMI is greater than to decrease until the 65 years old for both sexes but between 65 and 74 years old, the female sex reverses this trend.

	15-24		25-34		35-44		45-54		55-64		65-74		75+	
	$\bar{x}$	*Sd*	$\bar{x}$	*Sd*	$\bar{x}$	*Sd*	$\bar{x}$	*Sd*	$\bar{x}$	*Sd*	$\bar{x}$	*Sd*	$\bar{x}$	*Sd*
Male														
$\beta_{j,\cdot}$	7.5	7.6	6.7	7.3	11.9	9.3	15.8	9.4	13.8	8.0	19.8	10.6	26.4	10.3
$\phi_{j,\cdot}$	71.5	15.7	65.2	15.2	69.7	18.4	65.3	18.6	62.1	12.4	62.1	19.5	54.8	17.2
$\alpha_{j,\cdot}$	20.6	8.3	27.8	8.0	17.8	9.9	18.0	9.9	24.5	8.2	17.7	11.0	18.6	10.4
Female														
$\beta_{j,\cdot}$	19.9	8.4	9.1	8.6	9.8	8.7	11.0	10.7	13.3	11.4	20.4	11.3	10.2	9.6
$\phi_{j,\cdot}$	57.7	14.1	67.9	16.9	70.5	17.1	69.8	18.8	67.8	19.9	60.8	17.9	68.9	16.9
$\alpha_{j,\cdot}$	23.4	8.5	22.4	9.3	18.6	9.3	18.7	10.2	18.8	10.5	19.1	9.4	21.0	8.7

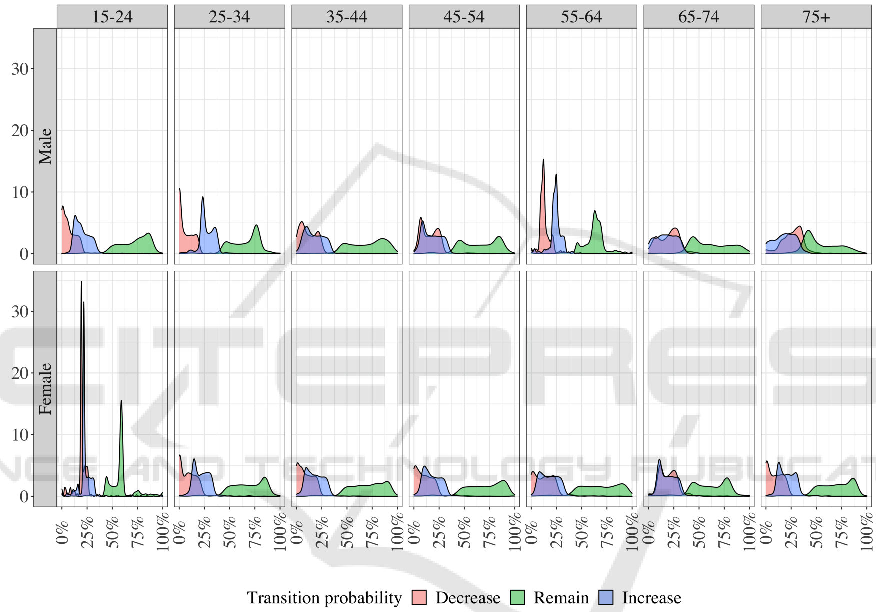

Figure 2: Distribution of the transition probabilities obtained for the fitting period (2003-2010) by age range and sex. The transition probabilities are ever in the same order; in decreasing order, they are Remain, Increase, Decrease; but via look-up analysis, we can suspect that this particular population tends to increase their BMI due to remaining a status just implies non-changes when the study is at short-term.

the low error in the fitting period, the variability of the transition probabilities distribution can make them noisy and their interpretation might be carefully done. However, the level of disaggregation of the variables plays an important role in it. Thus, on the assumption that the transition probabilities are constant within a particular period and that they do depend on the age and sex of the person, we get that the population under study has a clear trend, in the long-term sense, to increase their BMI, but it seems to be some variables that are not included yet. Besides, note that the non-linearity of the mathematical model plays an important role since the non-integer variables cannot be solved. Also, the lack of information, especially the time windows for collecting the data, is an important problem that adds noise to the model when non-stationary variables are considered. Also, an important discussion must be done about the obtained results, where it is known that any long-term forecasting or extrapolation may be wrong when the period gets longer. In this case, we can see an excellent estimation at first glance in the fitting period, but the forecasting suffers a change of trend, especially for sex female, in the validation period. Anyway, the transition probabilities distribution shows that there are differences not just between sexes, but between age ranges as well. In general terms, two age ranges that have a remarkable change of shape, at 55-64 years

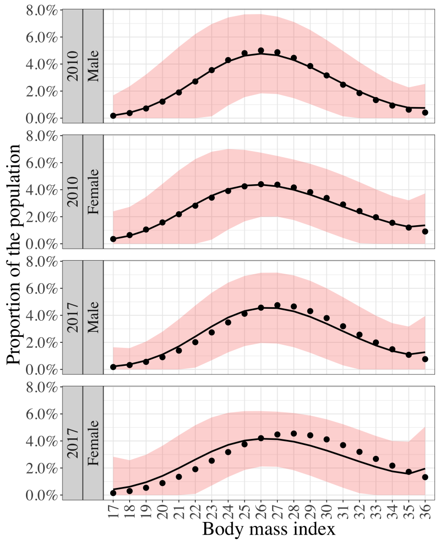

Figure 3: Comparison of the population obtained at the end of the fitting period (2010) with a 5.11% of total absolute error, and that one obtained from the forward simulation at the end of the validation period (2017) with a 10.27% total absolute error, by sex. The figure shows the known data as points and the fitted values are shown as lines, whereas the shaded area around them represents the 95% confidence interval.

old for sex men and at 15-24 years old for sex female. Those age ranges might be associated with hormonal changes in the body for the respective sex, such as the andropause for sex male (Tan and Pu, 2002) and the menstrual cycle for sex female (Van Hooff et al., 2004). Finally, we propose to explore the transition probabilities estimation by relaxing the assumption on the constant behavior according to the current BMI and to consider non-constant variables based on a different point of view as the Bayesian analysis.

ACKNOWLEDGEMENTS

The authors are grateful for partial support from the following sources: ANID Beca Magíster en el Extranjero, Becas Chile, Folio 73190041 (Javier Moraga-Correa) and Folio 73201112 (Luis Rojo-González), CONICYT-FONIS SA14ID0176 and RCUK-CONICYT Newton-Picarte MR /N026640/1 (Cristóbal Cuadrado), CMM-ANID AFB 170001 and CIMT-CORFO Cost Center 570111 (Jocelyn Dunstan), Universidad de Santiago de Chile, Proyecto DICYT 061817VP (Óscar C. Vásquez).

REFERENCES

Apovian, C. (2016). Obesity: definition, comorbidities, causes, and burden. *The American Journal of Managed Care*, 22(7 Suppl):s176–85.

Berrington de González, A., Hartge, P., Cerhan, J., et al. (2010). Body-mass index and mortality among 1.46 million white adults. *The New England Journal of Medicine*, 363(23):2211–2219.

Booth, H., Prevost, A., and Gulliford, M. (2012). Epidemiology of clinical body mass index recording in an obese population in primary care: a cohort study. *Journal of Public Health*, 35(1):67–74.

Departamento de Estadística e Información de Salud (2018). Defunciones. https://public.tableau.com/profile/deis4231#!/vizhome/Anuario_Defunciones/Defunciones. Online; accessed 3 oct 2020.

Detournay, B., Fagnani, F., Phillippo, M., et al. (2000). Obesity morbidity and health care costs in France: an analysis of the 1991-1992 medical care household survey. *International Journal of Obesity*, 24(2):151–155.

Fildes, A., Charlton, J., Rudisill, C., et al. (2015). Probability of an obese person attaining normal body weight: cohort study using electronic health records. *American Journal of Public Health*, 105(9):e54–e59.

Finkelstein, E., Fiebelkorn, I., and Wang, G. (2003). National medical spending attributable to overweight and obesity: How much, and who's paying? further evidence that overweight and obesity are contributing to the nation's health care bill at a growing rate. *Health Affairs*, 22(Suppl1):219–226.

Food and Agriculture Organization (2018). Panorama de la seguridad alimentaria y nutricional en américa latina y el caribe.

Frerichs, L., Araz, O., and Huang, T. (2013). Modeling social transmission dynamics of unhealthy behaviors for evaluating prevention and treatment interventions on childhood obesity. *PloS One*, 8(12):e82887.

Huxley, R., Mendis, S., Zheleznyakov, E., et al. (2010). Body mass index, waist circumference and waist: hip ratio as predictors of cardiovascular risk—a review of the literature. *European Journal of Clinical Nutrition*, 64(1):16–22.

Instituto Brasileiro de Geografía e Estadística, Brasil (2020). Pesquisa Nacional de Saúde - PNS. https://www.ibge.gov.br/estatisticas/sociais/justica-e-seguranca/9160-pesquisa-nacional-de-saude.html?edicao=9177&t=o-que-e. Online; accessed 3 oct 2020.

Instituto Nacional de Estadística (2017). Proyecciones de Población. https://www.ine.cl/estadisticas/sociales/demografia-y-vitales/proyecciones-de-poblacion. Online; accessed 3 oct 2020.

Laitinen, J., Power, C., and Järvelin, M. (2001). Family social class, maternal body mass index, childhood body mass index, and age at menarche as predictors of adult obesity. *The American Journal of Clinical Nutrition*, 74(3):287–294.

Lartey, S., Lei, S., Otahal, P., et al. (2020). Annual transition probabilities of overweight and obesity in older adults: Evidence from world health organization study on global ageing and adult health. *Social Science & Medicine*, 247:112821.

Levy, E., Lévy, P., Le, C., et al. (1995). The economic cost of obesity: the French situation. *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*, 19(11):788–792.

Manson, J., Skerrett, P., Greenland, P., et al. (2004). The escalating pandemics of obesity and sedentary lifestyle: a call to action for clinicians. *Archives of Internal Medicine*, 164(3):249–258.

Ministerio de Salud, Chile (2017). Base de datos. http://epi.minsal.cl/bases-de-datos/. Online; accessed 3 oct 2020.

Ministerio de Salud, Colombia (2015). Gestión del conocimiento y fuentes de información. https://www.minsalud.gov.co/salud/publica/epidemiologia/Paginas/gestion-del-conocimiento-y-fuentes-de-informacion.aspx. Online; accessed 3 oct 2020.

Mitchell, N., Catenacci, V., Wyatt, H., et al. (2011). Obesity: overview of an epidemic. *Psychiatric Clinics*, 34(4):717–732.

National Institutes of Health (1998). Clinical guidelines for the identification, evaluation, and treatment of overweight and obesity in adults - the evidence report. *Obesity Reviews*, 6(2):51S–209S.

NCD Risk Factor Collaboration (NCD-RisC) (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19.2 million participants. *The Lancet*, 387(10026):1377–1396.

NCD Risk Factor Collaboration (NCD-RisC) (2017). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128.9 million children, adolescents, and adults. *The Lancet*, 390(10113):2627–2642.

Ogden, C., Carroll, M., Curtin, L., et al. (2006). Prevalence of overweight and obesity in the United States, 1999-2004. *JAMA*, 295(13):1549–55.

Okorodudu, D., Jumean, M., Montori, V., et al. (2010). Diagnostic performance of body mass index to identify obesity as defined by body adiposity: a systematic review and meta-analysis. *International Journal of Obesity*, 34(5):791–799.

Olariu, E., Cadwell, K., Hancock, E., et al. (2017). Current recommendations on the estimation of transition probabilities in markov cohort models for use in health care decision-making: a targeted literature review. *ClinicoEconomics and Outcomes Research: CEOR*, 9:537–546.

Orpana, H., Tremblay, M., and Finès, P. (2006). *Trends in Weight Change Among Canadian Adults: Evidence from 1996/1997 to 2004/2005 National Population Health Survey*. Citeseer.

Penman, A. and Johnson, W. (2006). The changing shape of the body mass index distribution curve in the population: implications for public health policy to reduce the prevalence of adult obesity. *Preventing Chronic Disease*, 3(3).

Power, C., Lake, J., and Cole, T. (1997). Body mass index and height from childhood to adulthood in the 1958 british born cohort. *The American Journal of Clinical Nutrition*, 66(5):1094–1101.

Secretaría de Gobierno de Salud, Argentina (2019). 2° Encuesta Nacional de Nutrición y Salud ENNYS 2. Resumen ejecutivo. https://cesni-biblioteca.org/2-encuesta-nacional-de-nutricion-y-salud-ennys-2-resumen-ejecutivo/. Online; accessed 3 oct 2020.

Srinivasan, S., Bao, W., Wattigney, W., et al. (1996). Adolescent overweight is associated with adult overweight and related multiple cardiovascular risk factors: the bogalusa heart study. *Metabolism*, 45(2):235–240.

Sturm, R. (2002). The effects of obesity, smoking, and drinking on medical problems and costs. *Health Affairs*, 21(2):245–254.

Talukdar, D., Seenivasan, S., Cameron, A., et al. (2020). The association between national income and adult obesity prevalence: Empirical insights into temporal patterns and moderators of the association using 40 years of data across 147 countries. *PloS One*, 15(5):e0232236.

Tan, R. and Pu, S. (2002). Impact of obesity on hypogonadism in the andropause. *International Journal of Andrology*, 25(4):195–201.

Thomas, D., Weedermann, M., Fuemmeler, B., et al. (2014). Dynamic model predicting overweight, obesity, and extreme obesity prevalence trends. *Obesity*, 22(2):590–597.

Van de Kassteele, J., Hoogenveen, R., Engelfriet, P., et al. (2012). Estimating net transition probabilities from cross-sectional data with application to risk factors in chronic disease modeling. *Statistics in Medicine*, 31(6):533–543.

Van Hooff, M., Voorhorst, F., Kaptein, M., et al. (2004). Predictive value of menstrual cycle pattern, body mass index, hormone levels and polycystic ovaries at age 15 years for oligo-amenorrhoea at age 18 years. *Human Reproduction*, 19(2):383–392.

Wang, L., Denniston, M., Lee, S., et al. (2010). Long-term health and economic impact of preventing and reducing overweight and obesity in adolescence. *Journal of Adolescent Health*, 46(5):467–473.

Ward, Z. J., Long, M. W., Resch, S. C., et al. (2017). Simulation of growth trajectories of childhood obesity into adulthood. *The New England Journal of Medicine*, 377:2145–2153.

World Health Organization (2003). Controlling the global obesity epidemic. https://www.who.int/nutrition/topics/obesity/en/. Online; accessed 9 sep 2020.

World Health Organization (2016). Obesity and overweight. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight. Online; accessed 9 sep 2020.

Xue, H., Slivka, L., Igusa, T., et al. (2018). Applications of systems modelling in obesity research. *Obesity Reviews*, 19(9):1293–1308.