Item selection by an extended Latent Class model: An application to nursing homes evaluation

Bartolucci, Francesco and Giorgio E., Montanari and Pandolfi, Silvia

26 April 2012

Item selection by an extended Latent Class model: an application to nursing homes evaluation

Francesco Bartolucci, Giorgio E. Montanari, Silvia Pandolfi

26th April 2012

Abstract

The evaluation of nursing homes and the assessment of the quality of the health care provided to their patients are usually based on the administration of questionnaires made of a large number of polytomous items. In applications involving data collected by questionnaires of this type, the Latent Class (LC) model represents a useful tool for classifying subjects in homogenous groups. In this paper, we propose an algorithm for item selection, which is based on the LC model. The proposed algorithm is aimed at finding the smallest subset of items which provides an amount of information close to that of the initial set. The method sequentially eliminates the items that do not significantly change the classification of the subjects in the sample with respect to the classification based on the full set of items. The LC model, and then the item selection algorithm, may be also used with missing responses that are dealt with assuming a form of latent ignorability. The potentialities of the proposed approach are illustrated through an application to a nursing home dataset collected within the ULISSE project, which concerns the quality-of-life of elderly patients hosted in Italian nursing homes. The dataset presents several issues, such as missing responses and a very large number of items included in the questionnaire.

Keywords: Expectation-Maximization algorithm, Polytomous items, Quality-of-life, ULISSE project

1 Introduction

The evaluation of long-term care facilities is assuming a role of increasing relevance due to the rapid growth of demand for long-term care services for elderly people. The main cause is represented by the rapid aging of the population and also by the changes in the family structure and in the socio-economic context. Furthermore, the recent international debate on demographic aging has been focused on the effect of aging on the welfare and health care system in various countries (Galasso and Profeta, 2007; Breyer et al., 2010). In this context, measuring health care quality and comparing nursing home performance represent a challenging issue to assure the quality of services and to allocate resources efficiently. For this end, Kane et al. (2003), among others, propose quality-of-life measures using an index obtained through Factor Analysis. Phillips et al. (2007) deal with the construction of indicators which measure nursing home performance and propose the use of such indicators to rank these types of facility in a certain geographical area. The indicators currently used to evaluate nursing home performance are based on data coming from surveys which are periodically carried out by public institutions; see, among others, Hirdes et al. (1998) and Mor et al. (2003). In the United States, the nursing homes are compared by means of a set of quality of care indicators obtained by a standardized resident assessment instrument (Zimmerman, 2003). In particular, the Center for Medicare and Medicaid Services is engaged in building quality measures for nursing homes using descriptive statistics as indicators; these indicators are generally referred to the psycho-physics conditions of elderly people hosted in these facilities. Moreover, Grabowski et al. (2004) examine the relationship between Medicaid payment rates and quality of nursing home care, since Medicaid is the dominant payer of U.S. nursing home services.

Among other authors dealing with nursing home data, Gajewski et al. (2006) investigate the inter-rater reliability of the nursing home survey process by means of a Bayesian Latent Class analysis. Prado-Jean et al. (2011) develop and verify a tool to detect depression in elderly patients. The evaluation of nursing home performance is also studied in a longitudinal context by Bartolucci et al. (2009), who use a Latent Markov model to estimate the effect of nursing homes on the probability of transition between latent states representing different levels of the health status.

Issues related to population aging are particularly relevant in Italy, which is one of the European countries with the highest proportion of elderly people, and where this proportion is expected to increase over the next few decades (Kohler et al., 2002). This situation is putting a great pressure on the health care system, that in Italy is characterized by a high level of heterogeneity and fragmentation across the country. In this context, the ULISSE project (Lattanzio et al., 2010) was carried out by the Italian Ministry of Health jointly with the Italian Society of Gerontology and Geriatrics in order to obtain relevant data for health care planning. The purpose is to document the change in elderly patients' health status and the ability of the health care system to satisfy their needs. The dataset obtained from this project was collected by the administration of a questionnaire to patients hosted in a certain number of Italian nursing homes. The questionnaire is made of items about different aspects of the quality-of-life and health status of these subjects. In particular, we focus on 75 polytomous items, which have been chosen as clarified later. These items were administrated to a sample of 1744 patients. Many items are ordinal, with categories ordered according to increasing difficulty levels in accomplishing a certain task or severeness of the health condition. Moreover, given the complexity of the study, there are many missing responses.

Motivated by the data collected within the ULISSE project, in this paper we propose an algorithm that may be used for item selection when a large number of items is included in a questionnaire made of dichotomous, or in general polytomous, items. In this situation, the questionnaire administration may be lengthy and expensive. Due to tiring effects, using a large number of items may also induce the respondent to provide inaccurate responses. This is particularly relevant when the questionnaire is periodically administered. Therefore, methods are of interest which allow us to select the smallest subset of items providing an amount of information close to that of the full set. These methods may lead to a reduction of the costs of the data collection process and a better quality of data. Item selection is also important as a preliminary investigation that may be useful to reduce the amount of data, so that they may be more easily analyzed by complex statistical models.

The proposed algorithm is based on an extended version of the Latent Class (LC) model. The LC model (Lazarsfeld, 1950; Lazarsfeld and Henry, 1968; Goodman, 1974) has become an important tool of analysis of data collected by questionnaires made of polytomous items. As

is well known, the model relies on a discrete latent variable, which defines a certain number of latent classes, and assumes the independence of the responses to the items given this variable. Therefore, its use is justified when the items measure one or more latent traits, such as the quality-of-life or the tendency to a certain behavior, which are not directly observable. In geriatrics, the LC model has been used to measure mobility disability (Bandeen-Roche et al., 1997), to study behavioral syndromes in Alzheimer' patients (Moran et al., 2004), and to test the validity of certain physical frailty measures (Bandeen-Roche et al., 2006). Moreover, Lafortune et al. (2009) uses the LC analysis to model the heterogeneity in elderly individuals' health status.

It is important to recall that the LC model produces a *model-based classification* (Fraley and Raftery, 2002) of the observed subjects in different clusters, which correspond to the latent classes. Once the model is fitted, a subject is assigned to the latent class corresponding to the highest posterior probability (i.e., the conditional probability of the latent class given the observe data). This characteristic of the LC approach is exploited by the proposed method for item selection. In fact, the method assesses the relevance of an item on the basis of the number of subjects for whom the classification changes with or without this item. Therefore, this method starts from the classification of the subjects based on the full set of items; then, it sequentially removes the items which, among the existing ones, have the smallest impact on this classification. This procedure is stopped when the difference with respect to the initial classification is considered to be excessive. The proposed method is comparable to that for item selection proposed by Dean and Raftery (2010) in a similar context. Note, however, that the latter one evaluates the impact of removing an item through the Bayesian Information Criterion (BIC, Schwarz, 1978; Kass and Raftery, 1995), whereas our method is directly based on the impact in terms of classification of subjects.

As indicated above, the model on which our approach relies is an extended version of the LC model. The extension is for dealing with missing responses, which are frequently encountered in data as those coming from the ULISSE project. Obviously, formulating non-realistic assumptions in dealing with missing responses may lead to a strong bias in the parameter estimates. The extended LC model we adopt is based on some kind of *latent ignorability* condition (Harel and Schafer, 2009), given the latent class. In particular, we assume that the indicator for the presence of the response is conditionally independent

of the "underlying" response variable (which is not always observable), given the latent variable. Then, missing responses are informative, in the sense that the latent classes are also characterized in terms of probability that a response is provided. For this model, we implemented an Expectation-Maximization algorithm (EM; Baum et al., 1970; Goodman, 1974; Dempster et al., 1977), which has a complexity comparable to that of the conventional LC model. On the other hand, the method set up by Dean and Raftery (2010) is not directly formulated for the case of missing responses.

In dealing with missing responses, a more common assumption is that data are *missing at random* (MAR; Rubin, 1976; Little and Rubin, 2002). The basic idea is that the probability that a response variable is observed only depends on the values of those other variables which have been observed (Lu and Copas, 2004). Under this assumption, the missingness mechanism does not depend on the unobserved data (given the observed data) and ignoring the information in the missing-data indicator is generally appropriate. Computational methods for handling missing data under this assumption have been developed using, for example, the EM algorithm; see, among others, Tanner (1996), Schafer (1997), Kenward and Molenberghs (1998) and Little and Rubin (2002).

Among the authors dealing with the problem of missing responses in connection with latent variable models, we consider Muthén et al. (1987), who deal with latent variable structural equation models, and O'Muircheartaigh and Moustaki (1999) who consider the problem of treating item non-response in the analysis of attitude scales. Moreover, Reboussin et al. (2002) develop an LC model for the analysis of longitudinal binary health outcomes under the hypothesis that the data are MAR. Among the other authors who employ an LC structure to jointly describe the pattern of missingness and the outcome of interest, we also refer to Roy (2003) and Lin et al. (2004).

The application of the proposed item selection method to the data coming from the ULISSE project leads to a strong reduction of the number of items. In particular, we found a subset of around one third of items that has a degree of informativeness close to that of the full set of items. Moreover, this application allows us to compare our approach to missingness with the more common approach based on the MAR assumption. We found evidence that missing responses cannot be ignored and, also considering the sensitivity of some items, we conclude that our approach is then preferable.

The reminder of this paper is structured as follows. The ULISSE dataset is described in the following section. In Section 3 we illustrate the extended LC model with missing responses together with a discussion of the missingness assumptions we formulate. In Section 4 we describe maximum likelihood estimation of the proposed model, via the EM algorithm. In Section 5 we illustrate the proposed item selection procedure based on measuring the impact on the classification of the sample units. In Section 6 we present the results of the application of the proposed approach to the ULISSE dataset. Finally, Section 7 provides main conclusions about the proposed approach.

The methods proposed in this paper have been implemented in a series of MATLAB functions that we make available to the reader upon request.

2 The ULISSE dataset

The ULISSE project ("Un Link Informatico sui Servizi Sanitari Esistenti per l'Anziano" - "a computerized network on heath care services for older people") is aimed at describing the health status of elderly patients who currently receive health care assistance in Italy (Lattanzio et al., 2010). The project was carried out by the Italian Ministry of Health jointly with the Italian Society of Gerontology and Geriatrics. The main aim of the study is to improve the knowledge of the characteristics and the quality of health care services provided to elderly patients in Italy. The study considers three different levels of health care assistance: that provided by acute care facilities, that provided by nursing homes, and that provided at home. Overall, 23 acute geriatric or internal medicine hospital units, 31 nursing homes, and 11 home care services participated in the project. In the analysis here presented, we consider only the data collected in nursing homes.

The ULISSE project is based on a longitudinal survey; in the nursing homes considered by the project, the total number of residents, or a maximum of 100 randomly selected residents for bigger nursing homes, were evaluated at admission and then re-evaluated at 6 and 12 months after the admission. Only long stay residents (i.e., permanently admitted to the nursing home) aged at least 65 years, were included in the study. This survey was carried out since 2004 through the repeated administration of a questionnaire which is filled up by the nursing assistant of each patient. For our analysis, we consider only the first interview,

which covers 1744 patients.

The detailed patients information were collected using the classification system VAOR (Valutazione dell'Anziano Ospite di Residenza) that represents the Italian version of the interRAI Minimum Data Set (MDS) for nursing home care (Morris et al., 1991; Hawes et al., 1997). The questionnaire covers several aspects of the health status of the elderly patients. From the original questionnaire, we singled out 75 items, including only those actually related to health conditions. The discarded items are related to treatments and other aspects that are not relevant for our application.

The selected items are grouped into eight different sections of the questionnaire concerning:

1. Cognitive Conditions (CC);
2. Auditory and View Fields (AVF);
3. Humor and Behavioral Disorders (HBD);
4. Activities of Daily Living (ADL);
5. Incontinence (I);
6. Nutritional Field (NF);
7. Dental Disorders (DD);
8. Skin Conditions (SC).

The 75 items are polytomously-responded, with categories generally ordered according to increasing difficulty levels in accomplishing a certain task or severeness of the health condition. The complete list of items, with the corresponding number of response categories, is reported in Appendix. Table 1 also shows the average of the percentage of missing values computed with respect to the items composing each section of the questionnaire.

In this application, the LC model allows us to classify subjects into latent classes corresponding to different degrees of severeness of elderly patients' health condition. This classification may have important implications on the system of financial support for long-term nursing homes and on the evaluation of their performance with respect to their ability to retain over time patients in the classes corresponding to better quality-of-life.

	section	% missing
1	(CC)	1.04
2	(AVF)	0.77
3	(HBD)	2.24
4	(ADL)	9.23
5	(I)	1.89
6	(NF)	5.82
7	(DD)	2.47
8	(SC)	6.75

Table 1: *Average percentage of missing values for each section of the questionnaire.*

Latent variable models for ordinal variables are also present in the literature, such as the Rating Scale model (Andersen, 1977; Andrich, 1978). In this model, the response probability is expressed as a function of one or more latent variables through a specific link function (e.g., the cumulative logit link). Given the ordinal nature of the items composing our questionnaire, this model could be also used in the present context. However we prefer to avoid such parametrization and to rely on a standard LC model. The main advantages are the simplicity of the approach and the need of a reduced the number of parametric assumptions.

3 The Latent Class model with missing responses

Let J denote the number of items in the questionnaire of interest and, for a sample of n respondents, let Y_{ij} denote the response variable for subject i and item j, with $i = 1, \ldots, n$ and $j = 1, \ldots, J$, which has l_j categories, from 0 to $l_j - 1$. Moreover, to take into account missing responses, we introduce the binary indicators R_{ij}: when subject i responds to item j, R_{ij} is equal to 1 and the label corresponding to the chosen response category is assigned to Y_{ij}; otherwise, R_{ij} is set equal to 0 and an arbitrary value is assigned to Y_{ij}. The value assigned to the missing response is arbitrary in the sense that it does not have any influence on the estimation results. We also introduce the notation Y_{ij}^* for the "underlying" response which is unobserved when $R_{ij} = 0$ and coincides with Y_{ij} when $R_{ij} = 1$. Finally, we let $\boldsymbol{Y}_i^* = (Y_{i1}^*, \ldots, Y_{iJ}^*)$, $\boldsymbol{R}_i = (R_{i1}, \ldots, R_{iJ})$, and $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{iJ})$.

In order to explain the dependence structure between the response variables, the LC model assumes the existence of a discrete latent variable U_i which has, *a priori*, the same

distribution for every subject i. This distribution is based on k support points, labeled from 1 to k. Each support point corresponds to a latent class in the population and has a specific weight (or *a priori* probability); these weights are denoted by $\pi_1, \ldots, \pi_k$. Moreover, the conditional probability that individual i in class u provides response y to item j is

$$\lambda_{j|u}(y) = p(Y_{ij} = y|U_i = u), \quad j = 1, \ldots, J, \; u = 1, \ldots, k, \; y = 0, \ldots, l_j - 1.$$

We also introduce the parameters

$$\eta_{j|u} = p(R_{ij} = 1|U_i = u), \quad j = 1, \ldots, J, \; u = 1, \ldots, k,$$

corresponding to the probabilities that a subject in latent class u provides the response to item j. Overall, the number of non-redundant parameters of the extended LC model is then

$$m = (k - 1) + k \sum_j l_j,$$

since it has kJ additional parameters with respect to the standard LC model, which rules out missing responses (Goodman, 1974). Therefore, under this extended version of the LC model, the latent classes are also characterized in terms of probability that a response is given. In order to make the model more parsimonious, it is also possible to assume certain constraints on the parameters $\eta_{j|u}$ (see also Section 4.3).

As in the standard LC model, we assume *local independence*. For the present model, this assumption is formulated by requiring that, for all i, the response variables in $\boldsymbol{Y}_i^*$ (i.e., $Y_{i1}^*, \ldots, Y_{iJ}^*$) and in $\boldsymbol{R}_i$ (i.e., $R_{i1}, \ldots, R_{iJ}$) are conditionally independent, given the latent variable U_i. Moreover, the condition of *latent ignorability* (Harel and Schafer, 2009) that we assume is formulated by requiring that the random vector $\boldsymbol{R}_i$ is conditionally independent of the random vector $\boldsymbol{Y}_i^*$ given U_i.

The assumption of local independence implies that

$$p(\boldsymbol{y}_i^*|u) = p(\boldsymbol{Y}_i^* = \boldsymbol{y}_i^*|U_i = u) = \prod_j \lambda_{j|u}(y_{ij})$$

and

$$p(\boldsymbol{r}_i|u) = p(\boldsymbol{R}_i = \boldsymbol{r}_i|U_i = u) = \prod_j \eta_{j|u}^{r_{ij}}(1-\eta_{j|u})^{1-r_{ij}},$$

where $\boldsymbol{y}_i^*$ denotes a realization of $\boldsymbol{Y}_i^*$ with elements y_{ij}^* and $\boldsymbol{r}_i$ denotes a realization of $\boldsymbol{R}_i$ with elements r_{ij}, where $j = 1, \ldots, J$. Then, on the basis of the assumption of latent ignorability and considering that Y_{ij} is a function of R_{ij} and Y_{ij}^*, we have that

$$\begin{aligned} p(\boldsymbol{r}_i, \boldsymbol{y}_i|u) &= p(\boldsymbol{R}_i = \boldsymbol{r}_i, \boldsymbol{Y}_i = \boldsymbol{y}_i|U_i = u) = p(\boldsymbol{r}_i|u)p(\boldsymbol{y}_i|u) = \\ &= \prod_j \eta_{j|u}^{r_{ij}}(1-\eta_{j|u})^{1-r_{ij}} \left[r_{ij}\ \lambda_{j|u}(y_{ij}) + (1-r_{ij})\right] = \\ &= \prod_j \left[\eta_{j|u}\lambda_{j|u}(y_{ij})\right]^{r_{ij}} (1-\eta_{j|u})^{1-r_{ij}}, \end{aligned}$$

with $\boldsymbol{y}_i$ denoting a realization of $\boldsymbol{Y}_i$ having elements y_{ij}, $j = 1, \ldots, J$. Finally, the *manifest probability* of the response pattern $(\boldsymbol{r}_i, \boldsymbol{y}_i)$ for subject i is

$$p(\boldsymbol{r}_i, \boldsymbol{y}_i) = \sum_u p(\boldsymbol{r}_i, \boldsymbol{y}_i|u)\pi_u.$$

Another quantity of interest is the posterior probability that a subject with observed response configuration $(\boldsymbol{r}_i, \boldsymbol{y}_i)$ belongs to latent class u. Using standard rules, this probability is equal to

$$p(u|\boldsymbol{r}_i, \boldsymbol{y}_i) = p(U_i = u|\boldsymbol{R}_i = \boldsymbol{r}_i, \boldsymbol{Y}_i = \boldsymbol{y}_i) = \frac{p(\boldsymbol{r}_i, \boldsymbol{y}_i|u)\,\pi_u}{p(\boldsymbol{r}_i, \boldsymbol{y}_i)}, \quad u = 1, \ldots, k. \tag{1}$$

These posterior probabilities are used to allocate subjects in the different latent classes, as will be clarified in the following.

As mentioned in Section 1, an alternative way to formulate an LC model in the present context is on the basis of the hypothesis that the responses are *missing at random* (MAR; Rubin, 1976; Little and Rubin, 2002). In particular, MAR assumption states that the probability of the observed missingness pattern, given the observed and the unobserved data, does not depend on the unobserved data. Therefore, the missing data are *ignorable* for likelihood-based inference, provided that the parameters of the missing data mechanism are separable from those of the model for the responses of interest. Thus, we can proceed without worrying about the model for missingness and a valid analysis can be done considering only

the observed data (Howell, 2008). The conditional probability of the response configuration, given the latent class, is simply

$$p(\boldsymbol{y}_i|u) = \prod_{j:r_{ij}=1} \lambda_{j|u}(y_{ij}),$$

and then

$$p(\boldsymbol{y}_i) = \sum_u p(\boldsymbol{y}_i|u)\pi_u. \tag{2}$$

Note that in this case the number of parameters is the same as the standard LC model, that is $(k-1) + k\sum_j (l_j - 1) = m - kJ$.

In the application to the ULISSE dataset, the MAR assumption seems to be unrealistic, since it is plausible that patients with severe health condition are less likely to respond to some items. Considering the above latent ignorability assumption, the latent classes are also characterized in terms of probability that a given response is provided. This can provide more information with respect to the standard MAR assumption. A more detailed discussion on the missing assumptions for the considered application, together with a comparison between the maximum likelihood estimates obtained under our assumption of latent ignorability and the MAR assumption, are illustrated in Section 6.3.

4 Maximum likelihood estimation

In the following, we outline maximum likelihood estimation of the proposed LC model via the Expectation-Maximization (EM; Baum and Petrie, 1966; Baum et al., 1970) algorithm and we deal with the choice of the starting values for this algorithm. We then briefly discuss maximum likelihood estimation under the MAR assumption.

Given the independence between the sample units, the log-likelihood function of the model formulated in Section 3 is

$$\ell(\boldsymbol{\theta}) = \sum_i \log p(\boldsymbol{r}_i, \boldsymbol{y}_i), \tag{3}$$

where $\boldsymbol{\theta}$ is a short-hand notation for all model parameters. In order to estimate these parameters, we maximize this function by the EM algorithm.

4.1 Expectation-Maximization algorithm

The EM algorithm is based on the *complete-data likelihood* that we could compute if we knew the value of the latent variable U_i for every subject i in the sample. This is equivalent to the knowledge of the latent class to which every subject belongs. We represent this information by the set of dummy variables z_{iu}, $i = 1, \ldots, n$, $u = 1, \ldots, k$, where z_{iu} is equal to 1 if subject i belongs to latent class u and to 0 otherwise. Then, we can write the complete-data log-likelihood as

$$\begin{aligned} \ell^*(\boldsymbol{\theta}) &= \sum_i \sum_u z_{iu} \log[p(\boldsymbol{r}_i, \boldsymbol{y}_i|u)\pi_u] = \\ &= \sum_i \sum_u z_{iu} \sum_j [r_{ij} \log \eta_{j|u} + (1 - r_{ij}) \log(1 - \eta_{j|u})] + \\ &+ \sum_i \sum_u z_{iu} \sum_j r_{ij} \log \lambda_{j|u}(y_{ij}) + \sum_u z_{+u} \log \pi_u, \end{aligned} \tag{4}$$

where $z_{+u} = \sum_i z_{iu}$ is the number of subjects in latent class u. Note that, given a configuration of dummy variables z_{iu}, we have an explicit solution for the maximum of $\ell^*(\boldsymbol{\theta})$ in terms of the model parameters. In particular, we have

$$\pi_u = \frac{z_{+u}}{n}, \quad u = 1, \ldots, k, \tag{5}$$

$$\lambda_{j|u}(y) = \frac{\sum_i I(y_{ij} = y) r_{ij} z_{iu}}{\sum_i r_{ij} z_{iu}}, \quad j = 1, \ldots, J,\ u = 1, \ldots, k,\ y = 0, \ldots, l_j - 1, \tag{6}$$

$$\eta_{j|u} = \frac{\sum_i r_{ij} z_{iu}}{z_{+u}}, \quad j = 1, \ldots, J,\ u = 1, \ldots, k, \tag{7}$$

where $I(\cdot)$ is the indicator function equal to 1 if its argument is true and to 0 otherwise.

In order to maximize the model log-likelihood, the EM algorithm alternates the following two steps until convergence, starting from an initial guess of the model parameters collected in $\boldsymbol{\theta}$:

- **E-step**: compute the conditional expected value of the complete-data log-likelihood $\ell^*(\boldsymbol{\theta})$ given the observed data and the current value of the parameters;
- **M-step**: update the model parameters by maximizing the expected value obtained at the E-step.

Both steps are usually simple to implement. In practice, the E-step consists of obtaining

the posterior expected value of every dummy variable z_{iu}, that is

$$\hat{z}_{iu} = p(u|\boldsymbol{r}_i, \boldsymbol{y}_i), \quad i = 1, \ldots, n, \; u = 1, \ldots, k,$$

which may be computed according to (1).

At the M-step we maximize the expected value of the complete-data log-likelihood, which is obtained by substituting every dummy variable z_{iu} in (4) with $\hat{z}_{iu}$, and in this way we update the parameter vector $\boldsymbol{\theta}$. For this maximization we exploit the formulae given by (5), (6), and (7), with $\hat{z}_{iu}$ instead of z_{iu}.

As mentioned in Section 3, the posterior probabilities $\hat{z}_{iu}$ may be used for the classification of the subjects in the sample, that is for the allocation of the subjects into the k latent classes. In particular, on the basis of the output of the EM algorithm, we assign each subject i to latent class u when $\hat{z}_{iu} = \hat{z}_i^*$, where $\hat{z}_i^*$ is the maximum of $\hat{z}_{i1}, \ldots, \hat{z}_{ik}$. For this reason, Magidson and Vermunt (2001) and Vermunt and Magidson (2002) refer to this kind of model as an *LC cluster model*.

We have also to mention that, whereas the LC model assumes that individuals are full members of one of the classes for the discrete latent variable, in statistical literature other latent classification techniques exist. For example, the Grade of Membership model (Woodbury et al., 1978; Manton and Woodbury, 1991) assumes that individuals can be partial members of more than one class of a continuous distribution of the latent variables (Erosheva, 2002, 2006). This model was also used to identify clinically meaningful health profiles in community-living elderly (Portrait et al., 1999; McNamee, 2004) and to develop disability profiles (Erosheva et al., 2007).

4.2 Initialization of the algorithm

A typical problem of the latent variable and finite mixture models is the multimodality of the likelihood. Obviously, in the presence of more local maxima, the EM algorithm converges to one of them, which is not ensured to be the global maximum. In this case, it is advisable to use a deterministic initialization strategy together with a random initialization strategy. The latter consists of repeatedly initializing the algorithm from randomly chosen starting values for the parameters. When more starting values are used, the final estimate is the

one corresponding to the largest likelihood value that has been found at convergence of the EM algorithm (Biernacki et al., 2003; Karlis and Xekalaki, 2003). This solution is not guaranteed to correspond to the global maximum; however, it is rather obvious that the chance of reaching the global maximum increases with the number of starting values that is adopted. The drawback of the likelihood multimodality is particularly severe in the LC model used in our analysis, due to the large number of items, and to the fact that these items generally have more than two response categories.

In our approach we adopt a deterministic initialization which is based on a hierarchical clustering procedure of the sample units. This procedure uses a suitable distance measure computed on the basis of the complete-data log-likelihood; see equation (4). Given two clusters of units, this distance is equal to the difference between the complete log-likelihood of the LC model computed when these clusters correspond to two separate latent classes and that computed when these clusters are joined in the same latent class. The hierarchical clustering algorithm starts from the case of one cluster for every unit, which corresponds to $k = n$, and then sequentially merges these clusters until defining a classification based on only one cluster. Then, the EM algorithm is started from the classification corresponding to a number of clusters equal to the number of latent classes k of the adopted LC model.

We also adopt a random initialization which is based on drawing each latent class weight π_u from a uniform distribution from 0 and 1 and then normalizing these random draws in a suitable way. Similarly, we randomly choose $\lambda_{j|u}(y)$ and $\eta_{j|u}$. We use a number of random initializations which increases with the number of latent classes, because the latter affects the number of parameters and then the expected number of local maxima; more details are given in Section 6.1 with specific reference to the application.

4.3 Estimation under missing at random assumption

The EM algorithm illustrated above may be also employed to fit the LC model under the MAR assumption. In this case, the model log-likelihood is simply

$$\ell_m(\boldsymbol{\phi}) = \sum_i \log p(\boldsymbol{y}_i) \tag{8}$$

where the manifest probability $p(\boldsymbol{y}_i)$ is computed as in (2). As clarified in Section 3, under the MAR assumption, the parameters to be estimated are the same as in the standard LC model, that is the class weights $\pi_1, \ldots, \pi_k$ and the conditional response probabilities $\lambda_{j|u}(y)$, which are included in the vector $\boldsymbol{\phi}$.

It is worth noting that the estimates of the parameters above, obtained under the MAR assumption, may be also obtained from the proposed LC model under the hypothesis that the probability of responding to an item, $\eta_{j|u}$, does not depend on the latent class. Formally, this hypothesis may be expressed as $H_0 : \eta_{j|u} = \eta_j$, $u = 1, \ldots, k$. When it holds, the log-likelihood of the extended LC model defined in (3) may be expressed as the sum of two components: the first is $\ell_m(\boldsymbol{\phi})$ defined in (8), whereas the second is given by $\ell_r(\boldsymbol{\eta}) = \sum_i \log p(\boldsymbol{r}_i)$, with

$$p(\boldsymbol{r}_i) = \prod_j \eta_j^{r_{ij}} (1 - \eta_j)^{1 - r_{ij}},$$

and where $\boldsymbol{\eta}$ is the vector of all parameters η_j.

In order to test the hypothesis H_0, we can rely on a Likelihood Ratio (LR) statistic, which consists of comparing the log-likelihood $\ell(\hat{\boldsymbol{\theta}})$ with $\ell(\hat{\boldsymbol{\theta}}_0)$, where $\hat{\boldsymbol{\theta}}$ is the unconstraint estimate of $\boldsymbol{\theta}$ under the proposed LC model, whereas $\hat{\boldsymbol{\theta}}_0$ is the constraint estimate under H_0. This statistic is equal to $D = -2[\ell(\hat{\boldsymbol{\theta}}_0) - \ell(\hat{\boldsymbol{\theta}})]$ and, under H_0, has an asymptotic χ^2 distribution with $(k-1)J$ degrees of freedom. Since the parameter estimates under H_0 are equal to those obtained under the MAR assumption, through this test we can have some evidence in favor of or against the MAR assumption.

5 Item selection procedure

Once the LC model is fitted with a fixed number of latent classes and the subjects are classified on the basis of the posterior probabilities $\hat{z}_{iu}$ as described above, the item selection algorithm we propose sequentially removes an item from the initial set until a certain stopping rule is satisfied. The purpose here is to select the optimal subset of items, in terms of classification accuracy, from the set of available ones. In particular, the algorithm sequentially eliminates those items that do not significantly change the classification of the subjects in the sample with respect to that based on the full set of items.

More precisely, let $\mathcal{A}$ denote the set of existing items at the end of a given step of this algorithm. At the next step, we compute, for every $j \in \mathcal{A}$, the proportion $F^{\mathcal{A}\setminus j}$ of subjects whose classification is modified when this item is removed with respect to the classification based on the full set of items. Then, the item corresponding to the minimum value of $F^{\mathcal{A}\setminus j}$, $j \in \mathcal{A}$, is removed. If the minimum value of $F^{\mathcal{A}\setminus j}$ is common to two or more items in $\mathcal{A}$, the item with the minimum value of the index $D^{\mathcal{A}\setminus j}$ is removed. This index corresponds to the following Kullback-Leibler distance

$$D^{\mathcal{A}\setminus j} = \sum_i \sum_u \hat{z}_{iu} \log \frac{\hat{z}_{iu}}{\hat{z}_{iu}^{\mathcal{A}\setminus j}},$$

where $\hat{z}_{iu}^{\mathcal{A}\setminus j}$ is the posterior probability computed according to (1) considering all the responses to the items in $\mathcal{A}$ apart from item j. Other distance measures between the posterior probabilities $\hat{z}_{iu}$ and $\hat{z}_{iu}^{\mathcal{A}\setminus j}$ could be used, such as a distance based on the square differences between these probabilities.

It is worth noting that our item selection procedure is based on the consideration that an item is not useful when it does not have a strong connection with the latent trait and, therefore, it has not a significant influence on the classification of the subjects. Moreover, at each step of the item selection algorithm described above, the values of $F^{\mathcal{A}\setminus j}$ and $D^{\mathcal{A}\setminus j}$ are computed on the basis of the parameter estimates obtained from the initial fitting of the extended LC model. These initial estimates are obtained on the basis of data coming from the administration of a questionnaire, which is approved and validated at international level, to a large sample of subjects. Therefore, we consider these estimates on the same footing as "true" parameter values. On the other hand, this criterion is in accordance with the use of these parameters for classifying a sample of subjects on the basis of consecutive administrations of the questionnaire. Furthermore, since the classification of the subjects is always based on the same parameter estimates, we always consider the same ordering for the latent classes, that leads the clusters to be always identified in the same way.

The approach we propose considers all the subjects in the sample when computing the indices $F^{\mathcal{A}\setminus j}$ and $D^{\mathcal{A}\setminus j}$. Alternatively, we can discard those subjects that have unusual sequences of responses or a large number of missing responses. For these subjects the classification may change, even if the deleted items have a low discriminating power. In order

to decide which subjects have to be discarded we use a threshold, such as 0.95, for the maximum value of the posterior probabilities $\hat{z}_{iu}$. The subjects having a maximum posterior probability lower than this threshold may be considered as having an unusual behavior in relation to the classification procedure and then can be ignored during the item selection algorithm.

As already mentioned in Section 1, an alternative item selection procedure is proposed by Dean and Raftery (2010). In particular, these authors propose a method in which the importance of an item is assessed by comparing two models, given the items already selected. In one model the item provides information about cluster allocation beyond that contained in the already selected items, and in the other model it does not. The two models are compared via the Bayesian Information Criterion (BIC; Schwarz, 1978), seen as an approximation of the Bayes Factor. On the other hand, in the item selection strategy we propose, we explicitly consider a criterion which takes into account the classification of the subjects that arises from a selected subset of items. From this perspective, our method seems more sensible when, as in the application to the ULISSE dataset, the questionnaire is specifically tailored to classify a sample of subjects on the basis of some latent characteristic. Moreover, differently from the strategy of Dean and Raftery (2010), our strategy is explicitly formulated so as to account for missing responses.

6 Application to the ULISSE dataset

In this section, we illustrate the results obtained from the application of the proposed approach to the ULISSE dataset. We first show the maximum likelihood estimates obtained from the initialization strategy of the EM algorithm described in Section 4.2. Then, we deal with the choice of the number of latent classes (k) of the adopted LC model and we report a comparison between the missing assumptions applied to the data at hand. Finally, we illustrate the approach introduced in Section 5 to select a reduced set of items.

6.1 Choice of the number of latent classes

With reference to the analyzed dataset, we fitted the extended LC model described in Section 3 for a number of latent classes k from 1 to 10. For each k, we initialized the EM

algorithm as described in Section 4.2 from both a large number of random starting values and from the output of the hierarchical clustering procedure. In particular, we used $2^{(k+1)}$ random initializations for k greater than 1. The maximum value of the log-likelihood obtained under the deterministic and the random strategy is reported in Table 2. Obviously, for each value of k we take the estimate corresponding to the highest log-likelihood at convergence of the EM algorithm.

	maximum log-likelihood	
k	deterministic initialization	random initialization
1	**-118,645.34**	**-118,645.34**
2	**-106,489.45**	**-106,489.45**
3	-102,788.03	**-102,647.50**
4	-99,512.87	**-99,367.87**
5	**-97,150.04**	-97,409.08
6	**-95,848.61**	-96,189.39
7	**-94,815.31**	-94,978.31
8	**-93,877.77**	-94,028.91
9	**-93,148.59**	-93,190.11
10	**-92,483.79**	-92,630.66

Table 2: *Maximum value of the log-likelihood obtained under the deterministic and the random initialization strategy. The highest log-likelihood value in each row is in boldface.*

On the basis of the results above, we select the number of latent classes adopting different criterion. However, since we are especially interested in the classification of the sample of the elderly patients on the basis of the questionnaire, we give more relevance to criteria which measure the quality of the classification. In particular, we mainly rely on the normalized entropy criterion (NEC; Celeux and Soromenho, 1996; Biernacki et al., 1999) which is based on the index defined as

$$\text{NEC} = \frac{-\sum_i \sum_u \hat{z}_{iu} \log \hat{z}_{iu}}{(\hat{\ell} - \hat{\ell}_1)} \tag{9}$$

for $k \geq 2$, where $\hat{\ell}$ is the maximum of the log-likelihood of the model of interest and $\hat{\ell}_1$ is the maximum log-likelihood for the one latent class model. For $k = 1$, we have NEC $= 1$ by convention. The optimal number of latent classes is that corresponding to the minimum value of NEC. This criterion is based on the value of the posterior probabilities $\hat{z}_{iu} = p(u|\boldsymbol{r}_i, \boldsymbol{y}_i)$, which are used to classify subjects into latent classes, and takes also into account the goodness-of-fit of the model, which is measured by the log-likelihood.

Further to NEC, we consider a criterion based on an alternative index to measure the

quality of the classification (Bartolucci et al., 2009), which is defined as

$$\mathrm{S} = \frac{\sum_i (\hat{z}_i^* - \hat{\pi}^*)}{n\,(1 - \hat{\pi}^*)}, \tag{10}$$

where $\hat{z}_i^*$ is the maximum, with respect to u, of the posterior probabilities $\hat{z}_{iu}$, and $\hat{\pi}^*$ is the maximum, with respect to u, of the maximum likelihood estimates of the class weights under the model with k classes. When all the probabilities $\hat{z}_i^*$ are close to 1, the classification provided by the model relies on well separated latent states. In this situation, index S will attain a value close to its maximum which is equal to 1. On the other hand, when classes are not well separated, the index S will attain a value close to 0.

Finally, as in common applications of the LC model, we consider information criteria, such as the Akaike Information Criterion (AIC; Akaike, 1973), which is based on the index

$$\mathrm{AIC} = -2\,\hat{\ell} + 2\,m, \tag{11}$$

and the Bayesian Information Criterion (BIC; Schwarz, 1978), based on the index

$$\mathrm{BIC} = -2\,\hat{\ell} + m\log(n). \tag{12}$$

In the above expressions, m stands for the number of free parameters of the model. The optimal number of latent classes is that corresponding to the minimum value of the indices AIC or BIC. Usually BIC is preferable to AIC, as the latter tends to overestimate the number of states; see Dias (2006) and Nylund et al. (2007) for simulation studies on the performance of these and other information criteria.

Table 3 displays the maximum log-likelihood of the proposed model together with the corresponding number of parameters and the value attained for the indices described above. We observe that the BIC index assumes the minimum value for $k = 8$, whereas AIC leads to choosing a larger number of classes. However, taking into account the purpose of this application, we chose $k = 5$ latent classes, as suggested by the criteria based on the indices S and NEC. Note that a large number of latent classes may have a negative impact on the interpretability of the latent classes, which, in our application, correspond to different degrees of impairment of the elderly people health status.

k	$\hat{\ell}$	m	AIC	BIC	S	NEC
1	-118,645.34	150	237,590.67	238,410.26	-	-
2	-106,489.45	495	213,968.90	216,673.54	0.9821	0.0028
3	-102,647.50	743	206,781.00	210,840.70	0.9821	0.0029
4	-99,367.87	991	200,717.74	206,132.50	0.9848	0.0025
5	-97,150.04	1,239	196,778.08	203,547.89	**0.9863**	**0.0020**
6	-95,848.61	1,487	194,671.21	202,796.09	0.9813	0.0028
7	-94,815.31	1,735	193,100.62	202,580.55	0.9808	0.0028
8	-93,877.77	1,983	191,721.54	**202,556.52**	0.9813	0.0029
9	-93,148.59	2,231	190,759.19	202,949.23	0.9827	0.0026
10	-92,483.79	2,479	**189,925.58**	203,470.68	0.9817	0.0029

Table 3: *Selection of the number of latent classes for the LC model. For each number of classes, $\hat{\ell}$ is the maximum log-likelihood of the model, m is the number of parameters, and indices* AIC*,* BIC*,* S *and* NEC *are defined in (9), (10), (11), and (12). In boldface are reported the quantities corresponding to the model selected by the indices.*

6.2 Parameter estimates

In this section we report the results of the estimation procedure obtained when considering the proposed LC model, with reference to the selected number of latent classes, $k = 5$. Since the items are categorical, with a different number of categories, we decided to summarize the estimated conditional response probabilities, $\hat{\lambda}_{j|u}(y)$, by a weighted average computed with respect to values equally spaced between 0 and 1, and with weights given by these estimated response probabilities. This quantity can be expressed as

$$\hat{\lambda}^*_{j|u} = \frac{1}{l_j - 1} \sum_y y \hat{\lambda}_{j|u}(y).$$

For each item, this corresponds to assigning a score between 0 and 1 to the different response categories, and then computing the average of the scores, given the corresponding response probabilities. In particular, a value of $\hat{\lambda}^*_{j|u}$ close to 0 corresponds to a low probability of suffering from a certain pathology, whereas a value close to 1 corresponds to a high probability of suffering from the same pathology. To summarize these results, we also computed $\hat{\bar{\lambda}}^*_{d|u}$ as the average of the values assumed by $\hat{\lambda}^*_{j|u}$ with respect to the items composing each section d of the questionnaire, with $d = 1, \ldots, 8$. Finally, in order to have a clearer interpretation of the results, we ordered the latent classes on the basis of the values of $\hat{\bar{\lambda}}^*_{d|u}$ assumed in the first section ($\hat{\bar{\lambda}}^*_{1|u}$), so that the first class may be interpreted as that of subjects with the best health conditions. For each latent class, Table 4 shows the values of $\hat{\bar{\lambda}}^*_{d|u}$ together with the

estimated class weights $\hat{\pi}_u$.

	d								
u	1 (CC)	2 (AVF)	3 (HBD)	4 (ADL)	5 (I)	6 (NF)	7 (DD)	8 (SC)	$\hat{\pi}_u$
1	0.115	0.134	0.089	0.118	0.275	0.065	0.205	0.022	0.281
2	0.170	0.183	0.108	0.585	0.627	0.103	0.216	0.076	0.206
3	0.573	0.487	0.248	0.546	0.783	0.179	0.139	0.099	0.030
4	0.624	0.389	0.234	0.310	0.644	0.086	0.206	0.023	0.182
5	0.693	0.544	0.163	0.781	0.885	0.168	0.214	0.129	0.301
$\max_u(\hat{\bar{\lambda}}^*_{d\|u}) - \min_u(\hat{\bar{\lambda}}^*_{d\|u})$	0.579	0.410	0.158	0.663	0.610	0.114	0.077	0.107	

Table 4: *Means of the estimated response probabilities, $\hat{\bar{\lambda}}^*_{d|u}$, for each latent class u and each section d of the questionnaire, together with the estimated weights $\hat{\pi}_u$, under the latent ignorability assumption.*

The average $\hat{\bar{\eta}}_{d|u}$ of the estimated probabilities, $\hat{\eta}_{j|u}$, of giving a response to the items composing each section d is reported in Table 5. We can see that the lowest estimated probabilities are around 0.5.

	d							
u	1 (CC)	2 (AVF)	3 (HBD)	4 (ADL)	5 (I)	6 (NF)	7 (DD)	8 (SC)
1	0.992	0.996	0.994	0.986	0.984	0.950	0.976	0.946
2	0.992	0.993	0.996	0.885	0.995	0.949	0.997	0.953
3	0.841	0.906	0.479	0.756	0.849	0.613	0.678	0.540
4	0.997	0.996	0.993	0.956	0.984	0.960	0.991	0.940
5	0.997	0.995	0.991	0.836	0.980	0.952	0.981	0.941

Table 5: *Means of the estimated probabilities of giving a response to an item, $\hat{\bar{\eta}}_{d|u}$, for each latent class u and each section d of the questionnaire, under the latent ignorability assumption.*

6.3 Comparison with the MAR assumption

The previous results have been obtained under the assumption of latent ignorability for the missing responses, illustrated in Section 3. In the following, we compare the results obtained under this assumption with those obtained under the MAR assumption. For this aim, in Table 6 (top panel) we report the means of the estimated response probabilities, which are denoted by $\tilde{\bar{\lambda}}^*_{d|u}$ and are computed as clarified in the previous section, together with the estimated class weights, $\tilde{\pi}_u$, obtained under the MAR assumption. In the same table

(bottom panel) we report the differences between these estimates and the corresponding estimates obtained under the latent ignorability assumption (Table 4).

	d								
u	1 (CC)	2 (AVF)	3 (HBD)	4 (ADL)	5 (I)	6 (NF)	7 (DD)	8 (SC)	$\tilde{\pi}_u$
1	0.099	0.123	0.077	0.110	0.262	0.063	0.204	0.020	0.253
2	0.133	0.166	0.079	0.555	0.592	0.095	0.221	0.077	0.177
3	0.561	0.359	0.236	0.233	0.561	0.085	0.200	0.022	0.171
4	0.574	0.373	0.215	0.682	0.828	0.157	0.226	0.078	0.199
5	0.740	0.642	0.145	0.788	0.906	0.159	0.195	0.147	0.199
1	-0.015	-0.011	-0.013	-0.008	-0.013	-0.002	-0.001	-0.002	-0.028
2	-0.037	-0.017	-0.029	-0.029	-0.035	-0.008	0.005	0.001	-0.029
3	-0.012	-0.128	-0.012	-0.313	-0.222	-0.095	0.061	-0.077	0.141
4	-0.050	-0.016	-0.020	0.372	0.184	0.071	0.020	0.056	0.017
5	0.047	0.098	-0.019	0.007	0.020	-0.008	-0.019	0.019	-0.102

Table 6: *Means of the estimated response probabilities, $\tilde{\bar{\lambda}}^*_{d|u}$, for each latent class u and each section d of the questionnaire, together with the estimated weights $\tilde{\pi}_u$, under the MAR assumption (top panel). Differences with respect to the corresponding estimates under the latent ignorability assumption (bottom panel).*

From the results in Table 6 we observe that the largest differences between the estimates obtained under the MAR assumption and under the latent ignorability assumption exist for sections ADL (Activity of Daily Living) and I (Incontinence). Note that large differences are especially observed for subjects in the third latent class, who have the lowest estimated probability of giving a response (see Table 5). Moreover, for this class, and for the last class, we observe the largest change in the estimated weight. Overall, ignoring the missing data mechanism leads to a significant difference in the estimation results, especially for those sections of the questionnaire in which the number of missing values is relevant, and for the classes that include patients with high probability of missing responses. Therefore, the MAR assumption seems to be restrictive for the analysis of the data at hand, whereas our assumption of latent ignorability seems more realistic.

The above conclusion is confirmed by the result of the test of the hypothesis $H_0 : \eta_{j|u} = \eta_j$, $u = 1, \ldots, k$, already illustrated in Section 4.3, under which we obtain the same parameter estimates as under the MAR assumption. With $k = 5$, the LR statistic for this hypothesis is equal to $D = 4,946.58$ with 300 degrees of freedom and then a p-value equal to 0. Therefore, we have a strong evidence against the hypothesis H_0 and then against the MAR assumption. On the other hand, it is rather difficult to consider missing responses as ignorable given the

sensitivity of certain items and the multidimensionality of the questionnaire.

Here, we also report the results of the estimation procedure when the LC model is applied to a reduced dataset in which all the patients with at least one missing response to an item are removed. In this way the number of patients is reduced from 1744 to 592. Table 7 shows the means of the estimated response probabilities, $\tilde{\bar{\lambda}}^*_{d|u}$, and the estimated class weights, $\tilde{\pi}_u$, together with the differences with respect to the results reported in Table 4. Even in this case, we observe that largest differences are detected for sections ADL and I of the questionnaire and for the third latent class. Overall, these differences are significant, confirming that missing responses are not ignorable.

	d								
u	1 (CC)	2 (AVF)	3 (HBD)	4 (ADL)	5 (I)	6 (NF)	7 (DD)	8 (SC)	$\tilde{\pi}_u$
1	0.079	0.098	0.051	0.080	0.240	0.049	0.207	0.016	0.304
2	0.112	0.152	0.075	0.396	0.459	0.083	0.215	0.060	0.191
3	0.354	0.224	0.227	0.152	0.311	0.093	0.202	0.016	0.187
4	0.599	0.429	0.187	0.720	0.809	0.167	0.259	0.065	0.140
5	0.669	0.422	0.246	0.351	0.693	0.107	0.225	0.033	0.179
1	-0.036	-0.035	-0.038	-0.039	-0.035	-0.016	0.002	-0.006	0.022
2	-0.058	-0.031	-0.033	-0.189	-0.169	-0.019	-0.001	-0.016	-0.015
3	-0.219	-0.263	-0.021	-0.394	-0.473	-0.086	0.063	-0.083	0.156
4	-0.025	0.040	-0.048	0.410	0.165	0.081	0.053	0.042	-0.042
5	-0.024	-0.122	0.082	-0.431	-0.192	-0.061	0.011	-0.096	-0.122

Table 7: *Means of the estimated response probabilities, $\tilde{\bar{\lambda}}^*_{d|u}$, for each latent class u and each section d of the questionnaire, together with the estimated weights $\tilde{\pi}_u$, under the reduced dataset (top panel). Differences with respect to the corresponding estimates under the latent ignorability assumption (bottom panel).*

Finally, we clarify that the comparison in Table 6 is based on the same number of latent classes ($k = 5$) selected under the extended LC model which incorporates the latent ignorability assumption. However, under the MAR assumption the same selection criteria adopted in this application may lead to selecting a different number of classes, but in this case a direct comparison in terms of parameter estimates would be infeasible. The same may happen when dealing with the reduced sample obtained by discarding all subjects with at least one missing response (see Table 7).

6.4 Item selection

On the basis of the estimates illustrated in Section 6.2, we performed the item selection algorithm described in Section 5. In this regard, two approaches may be adopted. The first uses the information on all subjects in the sample, whereas the second approach discards from the sample those subjects having an anomalous response pattern.

Under the first approach, the results of the item selection algorithm are reported in Table 8. In particular, with reference to the first 50 steps of the algorithm, the table indicates the item which is removed at each step, the section to which the item belongs, and the corresponding values of the indices $F^{\mathcal{A}\setminus j}$ and $D^{\mathcal{A}\setminus j}$.

step	item	section	$D^{\mathcal{A}\setminus j}$	$F^{\mathcal{A}\setminus j}$	step	item	section	$D^{\mathcal{A}\setminus j}$	$F^{\mathcal{A}\setminus j}$
1	65	NF	0.134	0.000	26	37	HBD	30.084	0.009
2	19	HBD	0.253	0.000	27	1	CC	30.319	0.009
3	63	NF	0.298	0.000	28	39	HBD	32.718	0.009
4	74	SC	0.548	0.000	29	35	HBD	36.347	0.010
5	25	HBD	0.762	0.000	30	66	NF	38.507	0.010
6	71	DD	1.274	0.000	31	55	ADL	42.424	0.011
7	31	HBD	1.360	0.001	32	73	SC	47.043	0.013
8	36	HBD	2.082	0.001	33	30	HBD	47.643	0.014
9	64	NF	2.880	0.001	34	47	ADL	57.895	0.015
10	24	HBD	3.274	0.001	35	49	ADL	70.899	0.015
11	22	HBD	4.106	0.002	36	27	HBD	77.196	0.018
12	75	SC	4.630	0.002	37	21	HBD	82.640	0.018
13	28	HBD	6.393	0.002	38	12	CC	75.490	0.020
14	60	I	5.656	0.002	39	69	DD	89.046	0.020
15	52	ADL	5.723	0.002	40	42	ADL	76.717	0.021
16	70	DD	7.845	0.002	41	54	ADL	92.204	0.021
17	23	HBD	8.766	0.003	42	34	HBD	100.836	0.023
18	18	AVF	9.678	0.003	43	72	DD	116.506	0.024
19	14	AVF	12.104	0.004	44	33	HBD	140.257	0.025
20	29	HBD	10.854	0.005	45	40	ADL	146.003	0.027
21	59	I	17.533	0.005	46	26	HBD	181.504	0.030
22	57	ADL	18.885	0.006	47	44	ADL	183.233	0.033
23	15	AVF	27.488	0.006	48	50	ADL	184.303	0.034
24	67	DD	19.878	0.007	49	13	CC	185.315	0.036
25	61	NF	30.078	0.007	50	53	ADL	196.592	0.036

Table 8: *Results of the item selection strategy, considering all the subjects in the sample.*

These results show that it is possible to remove up to 6 items without changing at all the classification of the subjects ($F^{\mathcal{A}\setminus j} = 0$ for the first six step). Moreover, we can drop up to 21 items changing the classification of only the 0.5% of the subjects. Overall, deleting 50 items, the classification changes for a percentage of subjects of around 4%. With respect to the sections of the questionnaire, we can see that most of the items that can be removed are

referred to the sections denoted by HBD (Humor and Behavioral Disorder), NF (Nutritional Field) and SC (Skin Condition), that represent the sections of the questionnaire that have a reduced discrimination power (see Table 4). Note that the values of $F^{\mathcal{A}\setminus j}$ do not increase monotonically, but this is not ensured from a mathematical point of view.

The second item selection approach only considers those subjects having a maximum value of the posterior probabilities, $\hat{z}_i^*$, greater than 0.95. When we apply this strategy to the ULISSE dataset, we observe that the 95.1% of the subjects may be kept in the sample. Considering this subset of subjects, Table 9 reports the values obtained by the two indices $F^{\mathcal{A}\setminus j}$ and $D^{\mathcal{A}\setminus j}$ for every step of the algorithm. By this strategy, the number of items that can be removed without changing the classification of the subjects increases from 6 to 25. Moreover, deleting 50 items, the percentage of subjects for whom the classification changes is lower than 2.5%.

step	item	section	$D^{\mathcal{A}\setminus j}$	$F^{\mathcal{A}\setminus j}$	step	item	section	$D^{\mathcal{A}\setminus j}$	$F^{\mathcal{A}\setminus j}$
1	31	HBD	0.044	0.000	26	52	ADL	8.902	0.001
2	63	NF	0.056	0.000	27	70	DD	9.455	0.001
3	19	HBD	0.064	0.000	28	25	HBD	10.445	0.001
4	65	NF	0.118	0.000	29	51	ADL	12.312	0.002
5	22	HBD	0.149	0.000	30	59	I	16.842	0.002
6	75	SC	0.175	0.000	31	62	NF	19.067	0.002
7	60	I	0.214	0.000	32	66	NF	20.772	0.002
8	37	HBD	0.287	0.000	33	1	CC	24.569	0.003
9	61	NF	0.373	0.000	34	50	ADL	27.189	0.005
10	14	AVF	0.481	0.000	35	34	HBD	31.070	0.005
11	30	HBD	0.528	0.000	36	16	AVF	37.561	0.007
12	64	NF	0.651	0.000	37	23	HBD	40.287	0.007
13	39	HBD	0.770	0.000	38	71	DD	41.656	0.007
14	28	HBD	0.936	0.000	39	45	ADL	40.270	0.008
15	57	ADL	1.179	0.000	40	21	HBD	42.904	0.009
16	67	DD	1.337	0.000	41	72	DD	54.959	0.009
17	29	HBD	1.631	0.000	42	54	ADL	58.471	0.010
18	74	SC	2.021	0.000	43	24	HBD	61.179	0.012
19	18	AVF	2.307	0.000	44	53	ADL	68.975	0.014
20	27	HBD	2.866	0.000	45	69	DD	86.810	0.015
21	40	ADL	3.780	0.000	46	36	HBD	99.006	0.016
22	38	HBD	4.606	0.000	47	55	ADL	112.675	0.017
23	33	HBD	5.309	0.000	48	5	CC	114.207	0.020
24	12	CC	6.623	0.000	49	13	CC	118.671	0.022
25	35	HBD	7.685	0.000	50	4	CC	132.679	0.023

Table 9: *Results of the item selection strategy, considering only those subjects in the sample having a maximum value of the posterior probabilities, $\hat{z}_i^*$, greater than 0.95.*

For both strategies, we also note that there is no a clear relation between the removed items and the corresponding number of missing responses.

7 Conclusions

In this paper, we propose an algorithm for item selection which is based on an extended version of the Latent Class (LC) model (Lazarsfeld, 1950; Lazarsfeld and Henry, 1968; Goodman, 1974). In particular, this algorithm is aimed at finding the smallest subset of items which provides an amount of information that is close to the original set of items in terms of classification of the subjects in homogenous clusters. Moreover, the extended version of the LC model we propose, and then the item selection algorithm, can be used in the presence of missing responses. This model relies on a form of *latent ignorability* assumption (Harel and Schafer, 2009), given the latent variable.

The method for item selection is of simple implementation. In particular, we implemented it in a series of MATLAB functions that we make available to the reader upon request. Moreover, in order to illustrate the potentiality of the proposed approach for large scale investigations, we report an application to a dataset collected within an Italian project, named ULISSE (Lattanzio et al., 2010), about the quality-of-life of elderly subjects hosted in a certain number of nursing homes.

The ULISSE dataset was collected by a very large number of polytomous items about several aspects of the health status of the patients and presents several missing responses. The extended LC model that we propose may be applied without discarding any record and then without loosing relevant information. In order to evaluate the validity of the latent ignorability assumption, we also illustrate a comparison between the estimation results obtained when assuming this condition and the results obtained when assuming the more standard *missing at random* (MAR) hypothesis (Rubin, 1976; Little and Rubin, 2002). For the data at hand, we found evidence against the MAR assumption, and then our approach, which considers missing responses not ignorable, seems more appropriate. On the other hand, this conclusion is in accordance with the particular topic dealt with by the questionnaire and its multidimensional structure.

Finally, consider that a large number of items may lead to a lengthy and expensive administration of the questionnaire and may induce the respondents to provide inaccurate responses. We show that the suggested algorithm for item selection, when applied to the ULISSE dataset, leads to a strong reduction of the number of items, without loosing relevant

information for the classification of the subjects. In particular, we found a subset of one third of items with a degree of informativeness close to that of the initial set. This implies clear advantages in terms of setting up a questionnaire which may be more easily administered to a sample of subjects, especially in a longitudinal context in which we have repeated measurements.

References

Akaike, H. (1973). Information theory and an extension of the Maximum Likelihood principle. In Petrov, B. and Csaki, F., editors, *Second International Symposium on Information Theory*, Budapest. Akademiai Kiado.

Andersen, E. (1977). Sufficient statistics and latent trait models. *Psychometrika*, 42:69–81.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43:561–573.

Bandeen-Roche, K., Miglioretti, D. L., Zeger, S. L., and Rathouz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, 92:1375–1386.

Bandeen-Roche, K., Xue, Q.-L., Ferrucci, L., Walston, J., Guralnik, J. M., Chaves, P., Zeger, S. L., and Fried, L. P. (2006). Phenotype of frailty: characterization in the women's health and aging studies. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 61:262–266.

Bartolucci, F., Lupparelli, M., and Montanari, G. E. (2009). Latent Markov model for longitudinal binary data: An application to the performance evaluation of nursing homes. *Annals of Applied Statistics*, 3:611–636.

Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state Markov chains. *The Annals of Mathematical Statistics*, 37:1554–1563.

Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (1970). A maximization technique

occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, 41:164–171.

Biernacki, C., Celeux, G., and Govaert, G. (1999). An improvement of the NEC criterion for assessing the number of clusters in a mixture model. *Non-Linear Analysis*, 20:267–272.

Biernacki, C., Celeux, G., and Govaert, G. (2003). Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics & Data Analysis*, 41:561–575.

Breyer, F., Costa-Font, J., and Felder, S. (2010). Ageing, health, and health care. *Oxford Review of Economic Policy*, 26:674–690.

Celeux, G. and Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13:195–212.

Dean, N. and Raftery, A. (2010). Latent class analysis variable selection. *Annals of the Institute of Statistical Mathematics*, 62:11–35.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum Likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39:1–38.

Dias (2006). Model selection for the binary latent class model: A Monte Carlo simulation. In *Data Science and Classification*, pages 91–99. Springer Berlin Heidelberg.

Erosheva, E. (2002). *Grade of membership and latent structure models with application to disability survey data*. PhD thesis, Carnegie Mellon University, Department of Statistics.

Erosheva, E., Fienberg, S., and Joutard, C. (2007). Describing disability through individual-level mixture models for multivariate binary data. *Annals of Applied Statistics*, 1:502–537.

Erosheva, E. A. (2006). Latent class representation of the grade of membership model. Technical report, University of Washington, Seattle.

Fraley, C. and Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97:611–631.

Gajewski, B., Thompson, S., Dunton, N., Becker, A., and Wrona, M. (2006). Inter-rater reliability of nursing home surveys: a Bayesian latent class approach. *Statistics in Medicine*, 25:325–344.

Galasso, V. and Profeta, P. (2007). How does ageing affect the welfare state? *European Journal of Political Economy*, 23:554–563.

Goodman, L. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika*, 61:215–231.

Grabowski, D., Angelelli, J., and Mor, V. (2004). Medicaid payment and risk-adjusted nursing home quality measures. *Health Affairs*, 23:243–252.

Harel, O. and Schafer, J. L. (2009). Partial and latent ignorability in missing-data problems. *Biometrika*, 96:37–50.

Hawes, C., Morris, J. N., Phillips, C. D., Fries, B. E., Murphy, K., and Mor, V. (1997). Development of the nursing home resident assessment instrument in the USA. *Age and Ageing*, 26:19–25.

Hirdes, J. P., Zimmerman, D., Hallman, K. G., and Soucie, P. S. (1998). Use of the MDS quality indicators to assess quality of care in institutional settings. *Canadian Journal for Quality in Health Care*, 14:5–11.

Howell, D. (2008). The treatment of missing data. In Outhwaite, W. and Turner, S., editors, *The Sage handbook of Social Science Methodology*, pages 208–224. London: Sage.

Kane, R. A., Kling, K. C., Bershadsky, B., Kane, R. L., Giles, K., Degenholtz, H. B., Liu, J., and Cutler, L. J. (2003). Quality of life measures for nursing home residents. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 58:240–248.

Karlis, D. and Xekalaki, E. (2003). Choosing initial values for the EM algorithm for finite mixtures. *Computational Statistics & Data Analysis*, 41:577–590.

Kass, R. and Raftery, A. (1995). Bayes factors. *Journal of the American Statitical Association*, 90:773–795.

Kenward, M. G. and Molenberghs, G. (1998). Likelihood based frequentist inference when data are missing at random. *Statistical Science*, 13:236–247.

Kohler, H., Billardi, F. C., and Ortega, J. (2002). The emergence of lowest-low fertility in Europe during the 1990s. *Population and Development review*, 28:641–680.

Lafortune, L., Beland, F., Bergman, H., and Ankri, J. (2009). Health status transitions in community-living elderly with complex care needs: a latent class approach. *BMC Geriatrics*, 9:6.

Lattanzio, F., Mussi, C., Scafato, E., Ruggiero, C., Dell'Aquila, G., Pedone, C., Mammarella, F., Galluzzo, L., Salvioli, G., Senin, U., Carbonin, P. U., Bernabei, R., and Cherubini, A. (2010). Health care for older people in Italy: The U.L.I.S.S.E. project (un link informatico sui servizi sanitari esistenti per l'anziano - a computerized network on health care services for older people). *J Nutr Health Aging*, 14:238–42.

Lazarsfeld, P. F. (1950). The logical and mathematical foundation of latent structure analysis. In S. A. Stouffer, L. Guttman, E. A. S., editor, *Measurement and Prediction*, New York. Princeton University Press.

Lazarsfeld, P. F. and Henry, N. W. (1968). *Latent Structure Analysis*. Houghton Mifflin, Boston.

Lin, H., McCulloch, C. E., and Rosenheck, R. A. (2004). Latent pattern mixture models for informative intermittent missing data in longitudinal studies. *Biometrics*, 60:295–305.

Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. Wiley Series in Probability and Statistics. Wiley, 2nd edition.

Lu, G. and Copas, J. B. (2004). Missing at random, likelihood ignorability and model completeness. *The Annals of Statistics*, 32:754–765.

Magidson, J. and Vermunt, J. K. (2001). Latent class factor and cluster models, bi-plots and related graphical displays. *Sociological Methodology*, 31:223–264.

Manton, K. G. and Woodbury, M. A. (1991). Grade of membership generalizations and aging research. *Experimental Aging Research*, 17:217–226.

McNamee, P. (2004). A comparison of the grade of membership measure with alternative health indicators in explaining costs for older people. *Health Economics*, 13:379–395.

Mor, V., Berg, K., Angelelli, J., Gifford, D., Morris, J., and Moore, T. (2003). The quality of quality measurement in U.S. nursing homes. *Gerontologist*, 43:37–46.

Moran, M., Walsh, C., Lynch, A., Coen, R. F., Coakley, D., and Lawlor, B. A. (2004). Syndromes of behavioural and psychological symptoms in mild Alzheimer's disease. *International Journal of Geriatric Psychiatry*, 19:359–364.

Morris, J., Hawes, C., Murphy, K., and et al. (1991). *Resident Assessment Instrument Training Manual and Resource Guide.* Eliot Press, Natick, MA.

Muthén, B., Kaplan, D., and Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, 52:431–462.

Nylund, K., Asparouhov, T., and Muthén, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, 14:535–569.

O'Muircheartaigh, C. and Moustaki, I. (1999). Symmetric pattern models: a latent variable approach to item non-response in attitude scales. *Journal of the Royal Statistical Society, Series A*, 162:177–194.

Phillips, C. D., Hawes, C., Lieberman, T., and Koren, M. J. (2007). Where should momma go? current nursing home performance measurement strategies and a less ambitious approach. *BMC Health Serv Res*, 7:93.

Portrait, F., Lindeboom, M., and Deeg, D. (1999). Health and mortality of the elderly: the grade of membership method, classification and determination. *Health Economics*, 8:441–457.

Prado-Jean, A., Couratier, P., Benissan-Tevi, L. A., Nubukpo, P., Druet-Cabanac, M., and Clement, J. P. (2011). Development and validation of an instrument to detect depression in nursing homes. Nursing homes short depression inventory (NH-SDI). *International Journal of Geriatric Psychiatry*, 26:853–859.

Reboussin, B. A., Miller, M. E., Lohman, K. K., and Have, T. R. T. (2002). Latent class models for longitudinal studies of the elderly with data missing at random. *Journal of the Royal Statistical Society, Series C*, 51:69–90.

Roy, J. (2003). Modeling longitudinal data with nonignorable dropouts using a latent dropout class model. *Biometrics*, 59:829–836.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63:581–592.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Chapman & Hall.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6:461–464.

Tanner, M. (1996). *Tools for statistical inference*. Springer-Verlag, New York, 3rd ed. edition.

Vermunt, J. K. and Magidson, J. (2002). Latent class cluster analysis. In Hagenaars, J. A. and McCutcheon, A. L., editors, *Applied latent class analysis*. Cambridge University Press, Cambridge, UK.

Woodbury, M. A., Clive, J., and Jr., A. G. (1978). Mathematical typology: A grade of membership technique for obtaining disease definition. *Computers and Biomedical Research*, 11:277–298.

Zimmerman, D. R. (2003). Improving nursing home quality of care through outcomes data: the MDS quality indicators. *International Journal of Geriatric Psychiatry*, 18:250–257.

Appendix

j	# cat.	item description
		Section CC
01	2	Short-term memory (0 = “recalls what recently happened (5 minutes)”, 1 = “does not recall”)
02	2	Long-term memory (0 = “keeps some past memories green”, 1 = “does not keep some past memories green”)
03	2	Memory status (0 = “recalls the actual season”, 1 = “does not recall the actual season”)
04	2	Memory status (0 = “recalls where is his room”, 1 = “does not recall where is his room”)
05	2	Memory status (0 = “recalls the names and faces of the staff”, 1 = “does not recall the names and faces of the staff”)
06	2	Memory status (0 = “recalls where he is”, 1 =“does not recall where he is”)
07	4	Decision about his daily activities (from 0 = “independent decisions” to 3 = “unable to decide”)
08	3	Easily sidetracked (from 0 = “problems absent” to 2 = “problems worsened in the last week”)
09	3	Altered perception or awareness of surrounding (from 0 = “problems absent” to 2 = “problems worsened in the last week”)
10	3	Disorganized speech (from 0 = “problems absent” to 2 = “problems worsened in the last week”)
11	3	Restlessness movements (from 0 = “problems absent” to 2 = “problems worsened in the last week”)
12	3	Lethargic spans (from 0 = “problems absent” to 2 = “problems worsened in the last week”)
13	3	Change in the cognitive conditions during the day (from 0 = “problems absent” to 2 = “problems worsened in the last week”)
		Section AVF
14	4	Hearing (from 0 =“no hearing impairment” to 3 = “severe hearing impairment”)
15	4	Ability to make itself understood (from 0 = “understood” to 3 = “seldom/never understood”)
16	3	Clear language (from 0 = “clear language” to 2 = “no language”)
17	4	Ability to understand others (from 0 = “understands” to 3 = “seldom/never understands”)
18	5	Sight in conditions of adequate lighting (from 0 = “no sight impairment” to 4 = “severe sight impairment”)
		Section HDB
19	3	Negative statements (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
20	3	Repetitive questions (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
21	3	Repetitive verbalizations (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
22	3	Persistent anger with himself or others (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
23	3	Self deprecation disesteem (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
24	3	Fears that are not real (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
25	3	To believe himself to be dying (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
26	3	To complain about his health (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
27	3	Repeated events anxiety (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
28	3	Unpleasant mood in morning (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
29	3	Insomnia/problems with sleep (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
30	3	Expressions of sad-faced (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
31	3	Easily tears (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
32	3	Repetitive movements (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
33	3	Abstention from activities of interest (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
34	3	Reduced local interactions (from 0 = “symptom not showed” to 2 = “symptom daily showed”)
35	4	To wander aimlessly (from 0 = “problem absent” to 3 = “problem daily encountered”)
36	4	Offensive language (from 0 = “problem absent” to 3 = “problem daily encountered”)
37	4	Physically aggressive (from 0 = “problem absent” to 3 = “problem daily encountered”)
38	4	Socially inappropriate behavior (from 0 = “problem absent” to 3 = “problem daily encountered”)
39	4	To refuse assistance (from 0 = “problem absent” to 3 = “problem daily encountered”)
		Section ADL
40	5	Moving to/from lying position (from 0 = “independent” to 4 = “totally dependent”)
41	5	Moving to/from bed, chair, wheelchair (from 0 = “independent” to 4 = “totally dependent”)
42	5	Walking between different points within the room (from 0 = “independent” to 4 = “totally dependent”)
43	5	Walking in the corridor (from 0 = “independent” to 4 = “totally dependent”)
44	5	Walking into the nursing home ward (from 0 = “independent” to 4 = “totally dependent”)
45	5	Walking outside the nursing home ward (from 0 = “independent” to 4 = “totally dependent”)
46	5	Dressing (from 0 = “independent” to 4 = “totally dependent”)
47	5	Eating (from 0 = “independent” to 4 = “totally dependent”)
48	5	Using the toilet room (from 0 = “independent” to 4 = “totally dependent”)
49	5	Personal hygiene (from 0 = “independent” to 4 = “totally dependent”)
50	5	Taking full-body bath/shower (from 0 = “independent” to 4 = “totally dependent”)
51	4	Balance problems (from 0 = “does not have balance problems” to 3 = “needs physical assistance”)
52	3	Mobility in the neck (0 = “no limitation”, 1 = “unilateral limitation”, 2 = “bilateral limitation”)
53	3	Mobility in the arm including shoulder or elbow (0 = “no limitation”, 1 = “unilateral limitation”, 2 = “bilateral limitation”)
54	3	Movements of the hand including wrist or finger (0 = “no limitation”, 1 = “unilateral limitation”, 2 = “bilateral limitation”)
55	3	Mobility in the leg and hip (0 = “no limitation”, 1 = “unilateral limitation”, 2 = “bilateral limitation”)
56	3	Mobility in the foot and ankle (0 = “no limitation”, 1 = “unilateral limitation”, 2 = “bilateral limitation”)
57	3	Other movements (0 = “no limitation”, 1 = “unilateral limitation”, 2 = “bilateral limitation”)
		Section I
58	5	Fecal incontinence (from 0 = “continence” to 4 = “incontinence”)
59	5	Urinary incontinence (from 0 = “continence” to 4 = “incontinence”)
60	2	Elimination of feces (0 = “adequate”, 1 = “not adequate”)
		Section NF
61	2	Chewing problem (0 = “no problem”, 1 = “problems”)
62	2	Swallowing problem (0 = “no problem”, 1 = “problems”)
63	2	Mouth pain (0 = “no problem”, 1 = “problems”)
64	2	Taste of many foods (0 = “does not complain”, 1 = “complains”)
65	2	Hungry (0 = “does not complain”, 1 = “complains”)
66	2	Food on his plate (0 = “does not leave it”, 1 = “leaves it”)
		Section DD
67	2	Debris present in mouth prior to going to bed at night (0 = “problem absent”, 1 = “problem present”)
68	2	Dentures/removable bridge (0 = “absent”, 1 = “present”)
69	2	Some/all natural teeth lost and does not have/does not use dentures (or partial plates) (0 = “problem absent”, 1 = “problem present”)
70	2	Broken, loose, or carious teeth (0 = “problem absent” 1 = “problem present”)
71	2	Inflamed gums, swollen or bleeding gums, oral abscesses, ulcers or rashes (0 = “problem absent”, 1 = “problem present”)
72	2	Dentures or removable bridge daily cleaned by resident or staff (0 = “absent”, 1 = “present”)
		Section SC
73	5	Pressure ulcer (from 0 = “no pressure ulcer” to 4 = “stage 4”)
74	5	Stasis ulcers (from 0 = “no pressure ulcer” to 4 = “stage 4”)
75	2	Resolved or cured ulcer (0 =“absent”, 1 = “present”)

Table 10: *Description of the full set of items.*