10

REVIEW AND DISCUSSION

In this thesis, several latent class models were described and applied examples of these models using datasets with varying participant characteristics were provided. This chapter will review the main findings of the previous chapters in the context of the current literature on latent class models; discuss some methodological issues and present future research directions.

Rationale and review

The acknowledgement of heterogeneity in study populations is not new. Within epidemiology, there has been a longstanding interest in grouping individuals into (predefined) normal versus "abnormal" categories. Well-known examples here include overweight categories, depression and alcoholism. Furthermore, clinicians categorise patients to simplify the complexity of reality in order to better guide their decision making processes (which person should be referred for further testing or treatment?). Categorisation therefore helps to distinguish those in immediate need of treatment or intervention from those who are not. Moreover, grouping patients based on their developmental trajectory of one or more risk factors or based on their risk-profiles also benefits the better understanding of the aetiology of diseases and the complex relationships between risk factors and disease (onset and progression). The categorisation of individuals (e.g. patients) is a central theme of this thesis; the focus is on the statistical derivation of (latent) subgroups, or classes. Chapters 4-9 described several applications of latent class models in a range of study populations. To a greater (for example chapter 7) or lesser (for example chapter 6) extent, each analysis revealed several (clinically) distinct subgroups with varying characteristics and the in-depth study of these distinct subgroups helped identify (small) groups of high-risk, or even low-risk individuals, who might, or might not, need special attention. For example, in chapter 7, distinct classes of type 2 diabetes patients with varying courses of diabetic retinopathy (DR) were identified. Several small classes of patients with unfavourable trajectories were found (e.g. showing progression in DR). These otherwise unobserved high-risk classes could have been masked if the data had been analysed using common statistical techniques such as mixed models [2] or generalised estimating equations [1].

Another cohort of type 2 diabetes patients was studied (cross-sectionally) in chapter 8. Possible heterogeneity in health care utilisation was explored, with three distinct classes being identified which were characterised by frequent home visits, low frequency contacts in surgery and high frequency contacts in surgery respectively. This identified heterogeneity in health care utilisation poses challenges for the development of structured disease management programs as disease management programs aim to structure and streamline health care for selected patient populations.

The abovementioned studies are good examples of how latent class models can be valuable as an exploratory tool in epidemiological research aiding the understanding of several aspects of type 2 diabetes.

To a lesser extent, the latent class models in chapter 6 revealed clinically distinct trajectories. Trajectories of vital exhaustion (VE) were studied and three latent classes were identified; a class showing no signs of vital exhaustion over time, a class showing stable, preclinical signs of VE over time and a small, relatively stable chronically vitally exhausted class. Differential (negative) cardiovascular health consequences later in life were not apparent after comparison of the three trajectories. These findings might be explained by the characteristics of the cohort used (i.e. the Amsterdam Growth and Health Longitudinal Study is a relatively homogenous cohort, as is the Spokane Heart Study, described in chapter 4 and latent class models might not have been the necessary statistical technique to apply in such situations). However, the main aim of the paper was to explore possible heterogeneity in trajectories of vital exhaustion (VE) in a healthy cohort, in part because previous research demonstrated this previously in a cohort of cardiac patients [173].

On the whole, latent class analyses can be a valuable statistical tool in many, but not all, situations. In the next sections of this chapter, the advantages, disadvantages, challenges and (methodological) issues are discussed that could help guide the choice for the application of latent class models in epidemiology.

Advantages of latent class models

Reviewing the previous chapters, several important advantages of latent class models can be identified. First, the most important advantage is that categorisation is not based on (subjectively defined) subgroups determined a-priori, but rather based on characteristics within the data and on statistical, more objective, criteria.

Another advantage is that the latent class approach (sometimes referred to as a person-centred approach) is particularly focused on relationships among individuals, aiming to categorise them in homogenous subgroups [35, 42]. These subgroups could represent multiple categories of patients with different (risk) profiles and subsequent health- or disease outcomes. This way of thinking moves away from the traditional way of summarising (large amounts of) patient-data into "*the average patient*". Instead, the person-centred approach acknowledges heterogeneity in the study population and summarises in a way *"multiple average patients"*.

Limitations of latent class models

Although the advantages are clearly appealing, there are also some important limitations and the models have been critically reviewed in the literature (see for example [39, 52, 262–265]). The first limitation concerns the complexity or flexibility of the models. Besides the potential computational heaviness and mathematical complexity of the models, there are numerous choices to make in the modelling process. Each choice influences the final number of classes and/or their characteristics. One of the most important examples concerns the longitudinal latent class models and has to do with the variance parameters in the model. If we fix within-class variance parameters to zero (i.e. by modelling LCGM), the final number of classes will often be larger compared to a model allowing freely estimated within-class variance parameters (i.e. in a LCGMM model). Moreover, general study design issues such as sample size and the number of measurement occasions have also been shown to influence the number and characteristics of identified classes in the final model [38–40, 266, 267]. In any setting, the consequences of each decision should be something that researchers need to keep in mind when conducting the latent class models: each decision could potentially influence the final outcome more than decisions in standard regression modelling would. These outcomes, of course, then drive the interpretations of the models and also subsequent implications.

Additional challenges for deciding on the final number of classes are encountered by the inconsistency of the model fit indices. Although simulation studies (see for example [17]) have shown that the BLRT is the most consistent model fit index, researchers still struggle with this problem in practice. The important issue on the whole is, that regarding all choices that can or need to be made in the modelling process, clear guidelines for building an appropriate latent class model are lacking, especially in epidemiology.

A third limitation has to do with the meaning of the latent classes. The extraction of the distinct classes (whether they are subgroups in latent class analysis or developmental trajectories in LCGM/LCGMM), however, is no proof for the true existence of multiple subpopulations (just as there is no proof of only one subpopulation when conducting for instance a mixed model analysis) [39, 52, 53, 268]. The final model is just a simplification of the complex reality (as are all statistical models), an issue which should be kept in mind when conducting latent class models.

Should latent class models be more widely adopted by epidemiologists?
If latent class models are applied correctly, in the appropriate settings and are interpreted with some caution, my opinion is that the advantages of latent class models outweigh the limitations and the models could be applied (more widely) in epidemiology. Besides keeping in mind the limitations described previously, there are some additional issues that should be considered if epidemiologists want to use these models. Important topics in this respect are 1) the research questions require the use of latent class models (heterogeneity within the data is hypothesised and this heterogeneity can be captured by *k* number of distinct latent classes); and 2) transparency and expertise is necessary.

1. The research questions require the use of latent class models (heterogeneity within the data is hypothesised)
The identification of distinct latent classes in epidemiological datasets, in particular in longitudinal datasets is often used as an exploratory tool. The aim is to unravel possible heterogeneity in trajectory shapes in a situation where the number of distinct subgroups and their characteristics are unknown. In some cases the identified trajectory shapes are further investigated by analysing predictors or consequences of them (as was done in the previous chapters of this thesis). Indeed, longitudinal latent class models (i.e. LCGM/LCGMM) are the statistical techniques of choice to answer research questions dealing with investigating and unravelling heterogeneity in trajectory shapes but, if investigation of potential predictors for the development over time in an outcome variable (e.g. investigating predictors of (heavy) tobacco smoking over time) is the primary aim, it is not always necessary to use latent class models. In such situations, this can be analysed by means of a mixed model analysis, which also incorporates a certain level of heterogeneity in development over time [2, 185]. In general, researchers should be aware that the application of latent class models (however appealing) are not necessarily the only solution to answer research questions dealing with (a certain level of) heterogeneity. If the research questions permits it, a mixed model analysis is quite capable of studying heterogeneity in development too.

Moreover, latent class model will always reveal a certain number of classes [265]. In some situations (it is probably impossible to recognise in which situations this is the case), the identified classes might not reflect clinically relevant or even clinically distinct subgroups. Researchers should therefore be aware that latent class models (again, however appealing) should only be applied on study populations where a certain degree of heterogeneity is expected. In other words, again, the research question should guide the statistical technique of choice, which is partly driven by previous literature and the characteristics of the study population. Looking at the applied examples in the previous chapters, clearly, the more heterogeneous study populations, such as the type 2 diabetes patients analysed in chapters 6 and 7 revealed clinically relevant as well as clinically distinct classes (a finding that was hypothesised by previous studies), whereas in the more homogeneous study populations of the Spokane Heart Study and the Amsterdam Growth and Health Longitudinal Study revealed subgroups showing comparable growth trajectories of BMI and vital exhaustion (VE) respectively, which differed mainly on baseline values and less on trajectory shape. In chapter 5, although the research question focussed on the identification of possible distinct trajectory shapes of VE, a mixed model analysis would also have been able to provide an answer on the research question, would we have focused only on *predictors* of VE-development (instead, we focused on *consequences* of VE-development, making a mixed model analysis not really possible).

2. Transparency and expertise

Latent class models are well known in the fields in which they were first developed in, and many accessible methodologically oriented papers [17, 20, 39, 51–53, 100, 262–264, 268, 269] have emerged describing the mathematical structure of the models including (bio) statistical issues in both real- and simulated data. These papers are slowly seeping through into the field of epidemiology [29, 120, 185, 270], helping to increase the confidence- and familiarity with these techniques. In this light, researchers conducting latent class models are encouraged to give insight in how the models were specified, because often the exact specification of the final model is unclear. This makes it difficult for researchers wanting to reproduce the identified classes or model in their own datasets. Especially when latent class applications are published in medical journals, most attention is paid to the interpretation and discussion of the classes for clinical practice and much less attention is paid to explaining the specification of the latent class model. Therefore, I encourage transparency, although this poses a possible dilemma on the researcher; often medical journals propose a (fairly) strict word limit for original contributions and the contents of the final manuscripts are a trade-off between the clinically orientated sections and the section describing the statistical analyses. What helps in this respect is a recent call for papers by the editorial board of the International Journal of Epidemiology [271]: the purpose of a new section in the journal, following a previous example of the Lancet [272], is to provide updates and/or reviews of recent statistical and epidemiological concepts and methods specifically from an educational perspective. Contributions published in these special issues can be used for (self-) educational purposes and can clarify complex issues in epidemiology. The editors state that they are eager to provoke discussion, clarify examples and misuse of methods. These special issues will pre-eminently be suitable for explaining the analysis steps of latent class models including the various ways of specifying latent class models could be a suitable topic in a future issue. These issues provide accessible information and learning material in order to increase the confidence- and familiarity with these techniques.

Unresolved issues

Latent class models are quickly evolving; (methodological) papers explaining, or criticising new or alternative ways of applying latent class models continue to be published and statisticians in particular continue to work on the expansion and combination of models (e.g. modelling in the Bayesian framework [273] and combining alternative models with latent class models [274]). There are too many developments to cover here, and some are even beyond the scope of this thesis. This section, therefore, is confined to some important issues that continue to dominate the literature and that I find particularly relevant for epidemiologists. These issues mainly concern methodological issues dealing with the application of the models in practice, but also some new developments of particular interest to the epidemiologist are discussed.

1: Handling of covariates: one-, two- or three-step approach?

Once the final number and characteristics of the latent classes are decided on, evaluating predictors and consequences of latent class membership are often the next step. Evaluating predictors and consequences of class membership profiles the types of individuals that belong to each of the classes [20] and may also serve as support for the validity of the classes [6, 19, 266, 267]. Moreover,

assessing (modifiable) predictors and long-term consequences of class membership provides valuable new insights for prevention and/or treatment [275]. There are many ways of handling these covariates, and although clear guidelines for how to incorporate covariates in latent class models and when which approach is most suitable are currently lacking, most researchers either use a one-step approach, or a two-step approach. The two-step approach, which was used in the previous chapters of this thesis, involves an approach where first latent class models are estimated to determine the final number of classes. In the next step, often done by conducting standard regression analyses, the characteristics of the classes are summarised and compared (in this second step either the predictors, or consequences or both of class membership are investigated). These steps can also be combined in a one-step approach, where these covariates are already included while estimating the (partly conditional) classes. Neither approach is right or wrong [269], but both approaches have been both favoured and criticised. For the two-step approach, the main advantage is that the covariates do not influence the class formation process, but because individuals are assigned to their most likely class, hereby ignoring uncertainty in the assignment process. The one-step approach does incorporate this uncertainty (and can do so in several ways [20]), but the covariates also influence the class formation process hereby clouding the interpretation of the latent classes. Moreover, one-step models do not always result in better model fit compared to two-step models [275] and the one-step models also have been shown to increase computation time and estimation problems are more likely to occur [20, 275], especially when there are many covariates.
The lack of clear guidelines, in combination with comparison studies [20, 275] showing differences in the outcomes (i.e. different approaches lead to different final number of classes and/or different class-characteristics) needs attention. Currently, most researchers make their decision based on practicality in combination with the focus of the research question. Research questions primarily of exploratory nature, dealing with the unravelling of heterogeneity (i.e. identifying the number, and characteristics of latent classes) are often answered by taking the two-step approach, where the classes are first estimated, individuals are classified in their most likely class and in a subsequent step, the classes are characterised to form the profiles of the classes. Although the one-step approach is generally the approach of choice in the fields of criminology and psychology [5], the two-step approach is appealing especially for applied researchers in epidemiology; the classes can be interpreted in a straightforward

manner and classification into the most-likely class (only if the posterior probabilities are high enough) results in an observed categorical variable that can be used in subsequent (regression) analyses.
Recently, Vermunt [70] described an improved three-step approach which has been implemented in the LatentGOLD statistical program [276]. This three-step approach takes into account the major disadvantages of both the one- and the two-step approach. In the improved approach, the latent class formation process in modelled first (i.e. the covariates are not clouding the interpretation of the formed classes). In the second step, the most likely class membership variable is created, which in the third step can be related to covariates (either predictors or consequences, or both), taking into account the measurement error (uncertainty) in class assignment. This approach, which was not implemented in Mplus until the end of September 2012, when version 7.0 was released, is a promising approach [70, 277]. In a sense, it uses the best of both approaches available previously and further research studying its value for epidemiology is justified.

2: Model building: the quest for the optimal number of classes and their characteristics
There continues to be a lively discussion on the model building process: there is no consensus about how and by which means to decide on the final number of classes [16–18], as described earlier. Two important issues are the main points of discussion: 1) making a decision based on the evaluation of established model fit indices, and/or 2) making a decision based on the degree of heterogeneity within a model (where allowing for less heterogeneity within a class often leads to an overestimation of the number of classes).
Inconsistency among model fit parameters is widely acknowledged and is one of the major limitations of latent class models. The general approach is therefore, is to compare multiple model fit indices with the clinical interpretation of the latent class solutions. On the degree of heterogeneity there is no consensus in the literature [29, 55, 120]. This part of the discussion focuses particularly on the differences between LCGM and LCGMM [5, 19, 29, 42, 266]. Some researchers argue that a choice for either a LCGM or a LCGMM should be made *before* analysing the data [29]. Others argue that this is not needed [6]. Some [18] even advise starting with an LCA also in longitudinal datasets (thus ignoring the correlated data), which can help to understand better what data there are; LCA is an ideal way to study heterogeneity in *patterns* [278] (which does not necessarily imply longitudinal *trajectories*), which can then

be studied further by LCGM and/or LCGMM. The discussion continues, but the question when to use LCGM or LCGMM remains. Currently, the general model building strategy is to start with the simplest model, i.e. a 1-class model and adding classes and comparing the models based on the criteria described in the previous chapters. The models in these first steps are often (variations of) LCGM, implying no within-class variation. Once the final model, or models, has/have been established, random intercepts and/or random slopes can then be added into one or more classes to assess if the model fit improves or if clinical need requires more within class flexibility. Therefore, the choice for the final model is always a trade-off between the outcome of the model fit indices (although they might not all point in the same direction) and clinical need. These criteria seem vague, and to a certain extent, they are, but it also shows clearly that researchers should be transparent about their reasons for choosing a final model (as described previously). Further research addressing these issues is needed as a correct decision on the number of classes is crucial for the clinical interpretation and generalisation of the identified subgroups.

3. Applications within epidemiology: focus on statistical programs and courses/ tutorials

In this thesis, Mplus [34] was used as the primary statistical software. Other software packages allow the analysis of similar models: LatentGOLD [276] and SAS [44] are two common alternatives. Both Mplus and LatentGOLD were initially developed to accommodate the more elaborate analysis of latent variables. Within epidemiology, these packages are not very common, probably because analyses with latent variables are uncommon [185]. Learning a new statistical software package can be time consuming and therefore epidemiologists might be hesitant to do so, hereby potentially hindering a widespread usage of latent class models. Accessible books, tutorials and courses focussing primarily on epidemiological studies (e.g. using examples from epidemiology), could increase the familiarity with new statistical software, and hereby increase familiarity with the latent class models. Such tutorials and courses, however, are currently not available on a large scale. The developers of almost all statistical software do provide help through web lectures and/or discussion boards, hereby creating an accessible way of sharing information and bringing researchers into contact with each other to focus on applications of latent class models relevant to epidemiological research questions.

4: Applications in randomised controlled trials and personalised medicine
In epidemiology, latent class models have been primarily conducted to answer a variety of research questions concerning the unravelling of heterogeneity in study populations. Examples of such research questions in this thesis have been addressed using data from observational cohorts, whereas studies analysing data with latent class models from randomised controlled trials (RCT) unfortunately are not part of this thesis. Relatively recently, several methodologically oriented papers [279–284] emerged describing the potential of latent class models for RCTs, in particular paying attention to the analysis of treatment effects in the presence of non-compliance and assessing heterogeneity in treatment effects. These papers provide interesting research opportunities that could be explored further by epidemiologists also.

Potential 1: CACE-models
The major interest of an RCT is the estimation of the treatment- or intervention effect. These effects are commonly analysed using the intention-to-treat (ITT) principle. Here, average treatment effects are estimated, ignoring possible non-compliance. The underlying assumption of this principle is therefore that every individual randomised in the treatment condition actually received the treatment. This, of course, is rarely the case. An as-treated analysis takes into account non-compliance, but ignores that compliance status is chosen by the individuals and therefore not randomised. This non-randomisation often results in compliers and non-compliers having different characteristics, which causes problems when statistically comparing these groups.
CACE-models, or complier average causal effect models, tackle both issues described above. The treatment effect is estimated only for compliers, but also compares the compliers in the treatment group with the (potential) compliers in the control group. The major difficulty, however, is that the complier status in the control group is unknown and this is where latent class models have shown to have potential if subgroups of individuals based on complier-status are regarded as unobserved subpopulations, analogous to latent classes. Thus, these subpopulations can have their own model parameters, or treatment- or intervention effects. Jo and Muthén [279] demonstrate that this is an efficient and interpretable approach of CACE-modelling. Although these descriptions proved of great value for researchers conducting RCTs, only few applications in epidemiology exist.

Potential 2: RCTs with repeated measures

Muthén and colleagues further demonstrated a novel approach for the analysis of RCTs with repeated measures [280, 281, 285]. Latent class growth mixture modelling can be utilised to assess heterogeneity in intervention effects in longitudinal trials. This approach allows the researcher to analyse the impact of an intervention on otherwise unobserved subgroups characterised by different growth trajectories. Subsequently, the findings can be used to determine meaningful benefits from intervention and evaluates when these benefits are likely to appear. These findings then aid the future design of intervention studies by the possibility to select different interventions for individuals belonging to different subpopulations. In a sense, this approach is comparable to the traditional "subgroup analyses", often conducted in papers describing the overall effect of an intervention [286]. Where in the traditional approach the subgroup analyses are determined by a-priori categorisation of individuals in predefined subgroups (e.g. gender, age, socio-economic status), whereas categorisation with LCGMM is on the basis of the (otherwise unobserved) heterogeneity in trajectories.

Potential 3: Personalised medicine?

Similar to the analysis of intervention effects in RCTs studying prevention programs, latent class models can also aid the better understanding of which treatment works for whom and when by categorising patients based on characteristics within the data and not based on (subjectively defined) subgroups determined a-priori. This way of thinking can aid the progression of personalised medicine [287–289] which proposes more individualisation of healthcare utilisation (e.g. cancer treatment or nutritional advice based on genetic profiling) and moves beyond data regarding *average* treatment- or intervention effects.

Epilogue

The main aim of this thesis was to study the applicability of latent class models for epidemiology, and I placed much focus on the methodology of the models throughout the thesis. The studies presented in chapters 4-9, did, however, contribute to greater or lesser extent valuable new insights for their respective research fields also. Taken together, I demonstrated that the recognition and subsequent analysis of heterogeneity within a study population can be a valuable insight, and can instigate new research ideas, provided that the models are conducted with caution. A multidisciplinary team of clinicians, epidemiologists and biostatisticians will therefore lead to the most fruitful results. Therefore, I encourage the continuing interaction between clinicians, epidemiologists and biostatisticians; awareness of the potential of latent class models for epidemiology, including the correct use of them will be beneficial to both the evolvement of the techniques as well as the epidemiological research field.