Qualified research is a valuable method for gaining detailed insight into phenomena under investigation. One common approach to qualitative data analysis is to transcribe interviews or focus group discussions, code the data, look for patterns among the codes, and then identify larger categories and themes (cf. Lincoln & Guba, 1985; Miles & Huberman, 1994; Strauss & Corbin, 1990). This process of coding, categorizing, and theming is completed by an individual researcher or a research team, either manually or with computer-assisted qualitative data analysis software programs (Miles & Huberman, 1994). Following such an approach, frequent and meaningful patterns can emerge, but this relies on researchers’ abilities to identify the patterns, even when they are using software to support this activity. When the amount of data collected in such projects is large, this approach to qualitative data analysis can become cumbersome and unwieldy. In such cases, important patterns that are subtle or complex can be missed.

One potential way to address these limitations is to employ a method for automatically discovering associations from among the codes that have been assigned to the data using traditional qualitative approaches. Association rules emerge as “if-then” rules in the form $A \Rightarrow C$. Here, $A$ (antecedents) and $C$ (consequents) are mutually exclusive sets of codes (i.e., attributes of the data). A particular association rule indicates that when the codes within set $A$ are present in the data, so are the codes in set $C$. Although this does not necessarily reflect a causal relationship between $A$ and $C$, it does show that there is an association or relationship between the codes worthy of further investigation. Association rules are normally accompanied by two measures of importance: support (the likelihood that a particular data element contains any of the codes within $A$ and $C$) and confidence (the likelihood that when the set of codes in $A$ appear within the data, so do the set of codes in $C$) (Agrawal & Srikant, 1994). These measures are often used to rank the discovered rules.

The process for discovering association rules was devised for the purpose of detecting patterns within large quantitative databases of sales transactions (i.e., market basket analysis). The method involves a statistical machine learning approach for identifying the degree to which all possible sets of antecedents are associated with all possible sets of consequents. Efficient algorithms have been developed to discover the collection of association rules in a reasonable amount of time, even when the data set is very large (Agrawal, Imielinski, & Swami, 1993). Although a large body of research has emerged within the domain of association rule discovery, including the development of open-source software tools (cf. Weka, 2010), to our knowledge there has been no work on the application of this technique as a data analysis tool for qualitative research.

There are several challenges to the application of association rule discovery within a qualitative research framework. First, since the algorithms for extracting the candidate association rules require the data to be within a database format, it is necessary to give careful consideration to how the data are modeled. The challenge is to identify a meaningful unit of analysis (i.e., event, object, or activity) that can be translated into a “transaction” within a database table. Each such transaction is entered as a row in the table, specifying each of the codes assigned to the transaction.

A second challenge to using association rules is that the ability to find interesting patterns depends upon the size of the data set. There must be a sufficiently large number of transactions in the database for patterns to emerge. If the data set is small, evidence regarding interesting and potentially relevant patterns might be scarce. The approach works best with a large data set for which it is difficult for a researcher to detect interesting patterns by inspection or using other manual or computer-aided data analysis methods.

Third, when applying this approach in qualitative research, it is very important to ensure that the codes assigned to the data are consistent. Any inconsistencies or mistakes in the coding process can lead to erroneous rules being discovered or important rules being missed. Further, since the rules only show the associations between antecedent and consequent codes, as in all qualitative research, it is important to return to the data (e.g., original transcripts) to further explore and explain the association rule discoveries.

The primary benefit of using this approach in qualitative research is the ability to handle large data sets. Not only do association rules show patterns that may have been found via traditional qualitative data analysis techniques, they also identify complex and subtle patterns that are difficult to discover with traditional approaches due to “data overload”. As such, the ability to generate theory from large qualitative data sets is enhanced. The association rules that are discovered can guide the researcher in
exploring the data, seeking explanations for these relationships that would have been difficult to identify using traditional methods.

During the presentation, in addition to introducing the process of association rule discovery and the opportunities and challenges in using this approach to enhance qualitative research methods, a case study to illustrate the utility of this method will be discussed. A large qualitative research project was recently undertaken in which 42 semi-structured interviews were conducted with the presidents of community sport organizations within Canada regarding innovations that were undertaken by their organizations. The unit of analysis was defined as the individual innovations, of which 188 were identified by the participants. Each of the innovations was coded based on (a) the features of the organization and (b) the sport, as well as (c) the determinants, (d) the process, and (e) the types of the innovations. By applying association rule discovery, we were able to uncover a number of interesting patterns within the data that would not have been realized via traditional qualitative data analysis methods. The presentation will conclude with advice for qualitative researchers who wish to employ such methods in their own data analysis practices.