Estimating the Rate of Concussions in British Columbia Minor Hockey Using Community Volunteer Collected Data.

by

Jeremy D. Hamm

B.Sc., Simon Fraser University, 2002

A PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the Department
of
Statistics and Actuarial Science

SIMON FRASER UNIVERSITY
March 2004

APPROVAL

Name: Jeremy D. Hamm

Degree: Master of Science

Title of project: Estimating the Rate of Concussions in British Columbia Minor Hockey Using Community Volunteer Collected Data.

Examining Committee: Dr. Boxin Tang
Chair

Dr. Carl Schwarz
Senior Supervisor
Simon Fraser University

Dr. Randy Sitter
Simon Fraser University

Dr. David Goodman
Simon Fraser University

Dr. Joan Hu
External Examiner
Simon Fraser University

Date Approved:

Abstract

The rate of concussions has been examined in elite levels of ice hockey but has yet to be studied in community, youth hockey in British Columbia where they are also thought to occur. Due to the relative rarity of the concussion, a large number of games need to be observed in order to gain a reliable estimate. This can become very costly if hired people are used to collect this data. Hired people are also limited in the amount of follow-up needed to confirm each concussion. Therefore, a more thorough and cost effective method of data collection is needed in order to obtain reliable rate estimates.

This project assesses the use of community volunteers as a valid source of data collection while examining the effect on concussion rate due to player age and ability. Concussion rates are modeled using general estimating equations coupled with an adjusted AIC, used for quasi-likelihood techniques. While current results are inconclusive, a new study design is proposed which will be both cost-effective, while adjusting for the possibility that community volunteers under report the true number of concussions.

Acknowledgements

First and foremost, I would like to acknowledge my family: My mom for her endless support. When studying needed to be done, dinner was ready and laundry was taken care of. Mom's a beauty. My dad for his interest in my work while allowing me to be my own critic. Geoff for our endless banter regarding all things Simpson's and Monty Python related.

Secondly, my committee members. I would like to thank: Carl for his endless amount of help and always being available when he was in town. The speed at which solutions enter his mind is strangely eerie and never fails to reiterate the fact that the things I know are greatly exceeded by the things I do not know. Dr Dave Goodman for his funding and need for a statistician. There are many things that a statistician can count in this world but because of him, I was able to count something that was related to hockey while being involved in a Canuck crazed, Leaf hating environment. Randy Sitter for his guidance throughout in both matters related to the project and matters outside the project. If it were not for his suggestion to apply to grad school in statistics, I would currently have a whistle around my neck and be watching kids climb a rope. I feel I chose the right path.

I promised myself I wouldn't cry - 'sniff'.

I would like to send a tremendous thank you to Nori. She was there to support me through the entire process and was able to explain to me crazy ideas put forth by Carl such as 'topic sentences' and 'flow of ideas'. I mean, come on, who does that?

Thanks to Ian Williamson for doing the leg work to collect the data. Although only in possession of two legs, he can trick people into thinking he has more (like Bruce Lee).

I would like to thank all of my peers in grad school. Many times I have only needed to start talking out loud about a problem and it would peak someone's interest enough for them to spend time helping me out. When I first started using statistical software, I figured I would never be able to get it. But there were many people who were willing to give me advice in terms of programming, analysis and data manipulation. These are job skills that I learned a lot more quickly by sharing an office with 10 or so students.

Finally, in terms of funding, I would like to thank Randy for his role in my applying for an entrance scholarship; the Canadian Institutes of Health Research for funding the data collection and my time analyzing it; and Farouk Nathoo for recommending me for the job he was leaving at Children's Hospital.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

There is an inherent risk of injury when playing any sport. In ice hockey, one such injury is the concussion, whose level of risk is becoming of great interest to the sports medicine field. Recent scientific and anecdotal evidence has indicated that long term cognitive deficits can occur as a result of sustaining one (or many) concussions. Previous studies have shown that concussions do occur at the highest skill levels such as the National Hockey League (NHL) (Wennberg and Tator, 2003) and the Canadian Junior Hockey League (CJHL) (Goodman et al., 2001). At these elite levels, each team has a trainer and affiliated physician who are responsible for player injury management. Players often do not have this direct access to medical care at lower skill levels. Concussions are thought to occur at these lower levels and, with more than 500,000 minor (youth) hockey players currently registered with Hockey Canada (Hockey Canada web-site, 2003), this group is of specific concern. Yet the concussion rate for these young players is unknown.

At the present time, the Motor Behaviour Laboratory (MBL) researchers in the School of Kinesiology at Simon Fraser University are examining the issue of concussions as well as promoting awareness throughout the minor hockey community in British Columbia. Part of the program to increase awareness includes providing evidence that concussions are actually occurring at this level. By providing reliable estimates of the concussion rate in minor hockey, the MBL researchers can work towards reducing this rate while making hockey a safer and more enjoyable game. Along these lines, educational interventions designed to teach players concussion symptoms as well as injury management and return to play guidelines are currently under development. In addition, researchers are collecting data regarding the types of mechanisms associated with concussions so that players, coaches and parents can be made aware of high risk situations.

In our efforts to reduce the risk of concussion, this project has two steps. One step is to obtain estimates of the concussion rate at the minor hockey level. The second step is to design a method of obtaining these estimates that can be used in future seasons. These steps will enable us to determine if efforts to inform the community on the effects of concussion are accompanied with a reduction of the rate of concussion over time.

1.1 Study Objectives

In order to gain a precise estimate of the concussion rate, a large sample of games needs to be observed due to the relative rarity of concussions. To hire people to observe all games and report concussions would be extremely costly in terms of both money and time. Another option is to exploit a resource already in attendance at most games. This resource is hockey parents. Parents of players, as a whole, are a very supportive group of people. They watch the majority of their children's games (usually because players need to be driven to arenas) and are already very active in league politics and team fund raising. Their dedication to the sport is demonstrated by making sure their children are present at 5:00 am practices and traveling great distances to ensure that players can make out of town games that are scheduled shortly after the work day ends. Parents of players know each of their respective teams schedules and so would be able to report on games whose times have changed due to rescheduling. They also have the ability to follow-up suspected concussions since they are in constant contact with both players, parents and coaches. If we incorporate parent volunteers from teams to report on concussions directly to the researchers, information could be obtained with minimal cost. Providing this method is effective, it should reduce the cost of these types of studies greatly.

Four main concerns arise over the use of parents due to the fact that these people are community volunteers. First, are parents willing to participate. It may be that a parent attending his or her child's hockey game simply wants to watch the game or socialize with other parents and not be laden with tasks such as data collection. At least one volunteer from each team needs to participate such that enough data can be collected so concussion rate estimates can be made with respect to sub-populations within minor hockey. If the effect on concussion rates due to certain factors is to be examined, the teams that participate need to encompass all possible factor level combinations. Second, is the data of high quality? Even if they agree to participate, will they do an adequate job such that we can be confident reporting estimates back to both the hockey community or the academic community? The third concern is that previous literature has indicated that concussions are under-reported (Goodman et al., 2001). This may be because community volunteers are unable to detect concussions all of the time. Therefore, in this thesis, the focus is on '*head-incidents of concern*' which encompass all incidents that could result in potential concussions. Incidents of concern should compensate for any concussions that were not reported due to slight symptoms as well as provide further information with regards to mechanisms of injury. However, by expanding to potential as opposed to actual concussions, the subjectivity of a response is increased in that an incident of concern to one person may not be an incident of concern to another. Lastly, we would like to repeat this study in future seasons; consequently the data structure needs to be such that analysis becomes a simple cookbook approach that can be done almost instantaneously.

As a result, the objectives of this project are fourfold.

1. To gain an estimate of the rate of 'head-incidents of concern' in minor hockey.

2. To assess the ability of volunteers to report 'head-incidents of concern' so that unbiased and reliable estimates can be formed.

3. To be able to determine whether incident rates are affected by factors such as age division, skill level and association size.

4. To design the study such that it can be implemented as efficiently as possible

annually and that data analysis is a minor step in the process.

1.2 Project Outline

Chapter 2 of this project discusses in detail the data collection methods, how team eligibility for participation in this study was determined, and the best way to recruit participants from the eligible teams. It also describes the use of a 'gold standard' which was used to determine whether or not volunteers were a reliable source of data. Chapter 3 outlines the techniques used for data analysis. Models were fit to the data using General Estimating Equations which are effective for longitudinal and correlated data (Liang and Zeger, 1986). Estimates from competing models were averaged to account for model-to-model variation using the Akaike Information Criterion (AIC) adjusted for quasi-likelihood methods. The results are discussed in Chapter 4. Study improvements are recommended in Chapter 5 such that the current study can be applied to other areas in BC. The recommendations are aimed at expanding the study while maintaining a relatively low cost.

Chapter 2

Study Design

2.1 Minor Hockey

The British Columbia Amateur Hockey Association (BCAHA) is the governing body of minor hockey in the province. Everyone under the age of 19 who plays organized minor hockey in BC belongs to the BCAHA. This organization is divided up into ten *regions* (Figure 2.1). Each *region* is further divided into various numbers of *leagues* (For example, Figure 2.2 outlines the *leagues* that compose the region of the Pacific Coast Amateur Hockey Association (PCAHA)). *Leagues* are in turn divided into *associations*, each representing a city or town. For example, the Port Moody Minor Hockey Association is limited to residents of the city of Port Moody, while Ridge Meadows Minor Hockey Association combines the municipalities of Maple Ridge and Pitt Meadows. The terminology used by the BCAHA classifies itself, as well as every sub-organization (e.g. PCAHA) as an Association. To avoid confusion, and for the purposes of this paper, the term *association* is used to refer to organizations at the lowest level (i.e at the community level (e.g. Port Moody)). Each *Region*, *League* and *Association* has an executive board with a president and various members who make decisions for their respective organizations, keeping with BCAHA guidelines. These guidelines set the same player age divisions and skill categories for each Association. There are six divisions, each encompassing a two-year age span. Players can begin participating at the age of six and continue on until the year they turn eighteen.

The divisions are labeled tyke (ages 6-7), novice (ages 8-9), atom (ages 10-11), pee wee (ages 12-13), bantam (ages 14-15), midget (ages 16-17) and juvenile (ages 18-19). From the atom division onwards, players can compete to make teams representing skill categories A, B or C within each age division. 'A' teams are comprised of the highest skilled players while the lowest skilled players are on 'C' teams.

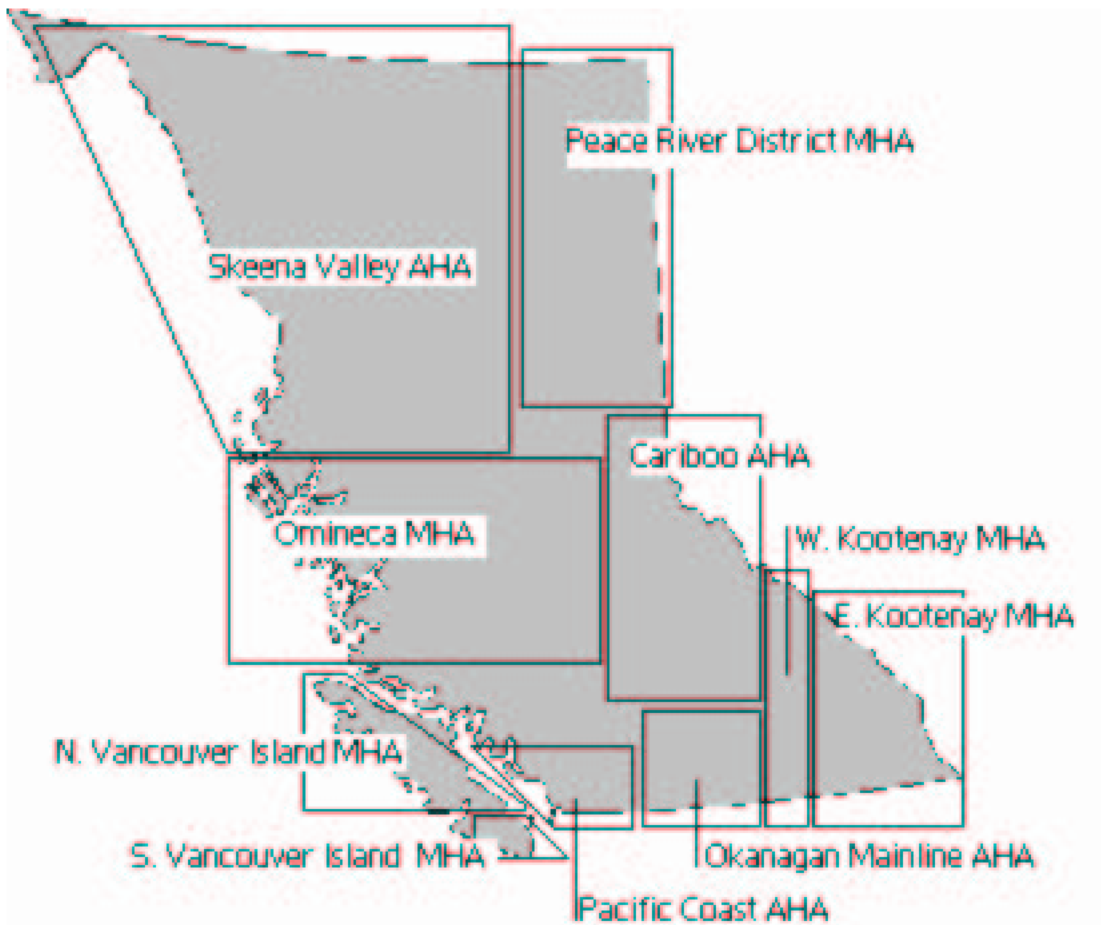


Figure 2.1: Minor hockey regions within the British Columbia Amateur Hockey Association.

The structure of the associations in terms of age and skill raises natural questions regarding the rate of incidents of concern. Questions such as do older (more experienced) players have a lower incident rate than younger (less experienced) players? Or will the increased size and strength that comes with increased age lead to an incident

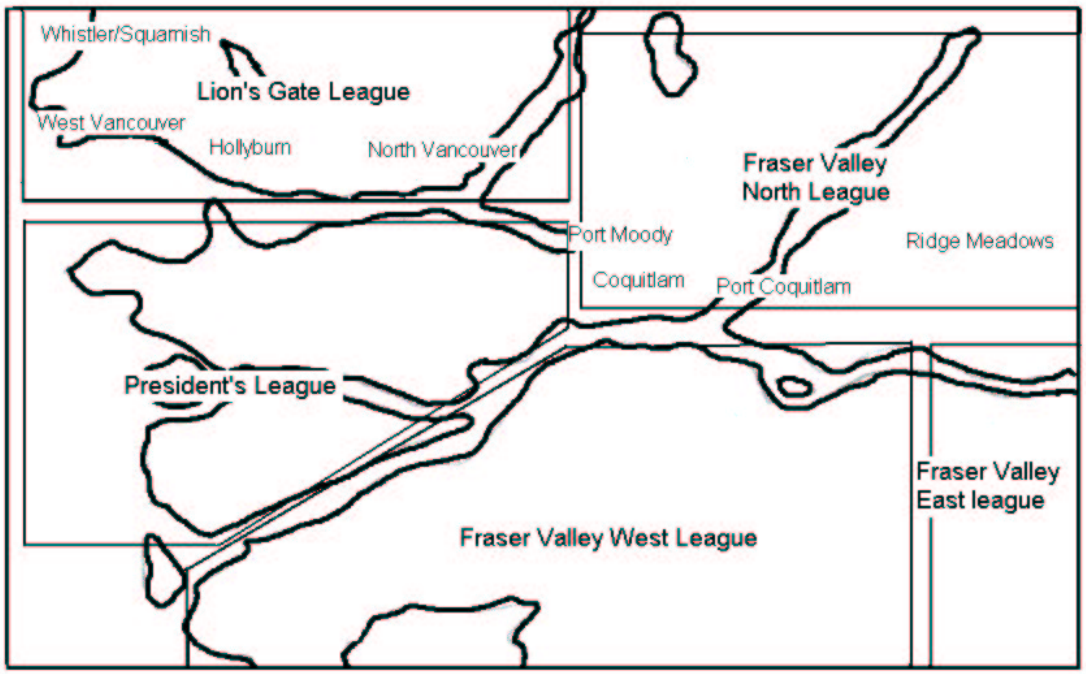


Figure 2.2: Leagues comprising the Pacific Coast Amateur Hockey Association.

rate higher than that observed among smaller, younger players? In terms of skill, is the rate higher for highly skilled players who move quicker and can body check harder? Or, does the high level of skill actually lower the incident rate due to players knowing how to properly receive a body check? While these questions are difficult to answer specifically, overall trends can be examined.

2.2 Study Factors

There were four factors in the study design: age division, skill category, association size and amount of researcher-volunteer contact.

Age division and skill category were included in order to examine how the structure of minor hockey in BC influences incident rates. However, not all age divisions were examined. Of the six, only three players (pee wee, bantam and midget) were observed because we expected them to have the highest incident risk. By age 18, many players drive themselves to games making volunteer recruiting much too difficult. The younger players were excluded to control cost. Association size and amount of researcher-volunteer contact were included as factors to examine the characteristics of volunteer participation.

Association size was composed of two levels, small and large. A small association was defined as having less than 13 teams and a large association had 13 or more teams. These definitions were arbitrary in nature, but allowed us to have equal numbers of small and large associations (four per level). The motivation behind this factor was that perhaps a smaller association acts as a more interactive community and may therefore, have a higher volunteer participation rate with more valid reports.

The factor of researcher-volunteer contact was designed to address the study's second objective -to assess the ability of volunteers to report incidents of concern so that unbiased and reliable estimates can be formed. There were two levels of researcher-volunteer contact: low and high. These two levels were attempts at determining the minimum anount of time and effort needed by the researchers to ensure that volunteers collected valid data. The volunteers in the low contact group (where each team was represented by a volunteer) were in contact with researchers only at the beginning, middle and end of the season in order to address questions or concerns. The volunteers in the high contact group were contacted by researchers every two months to address questions and concerns, as well as to ensure that things in general were going well. The question that remained was: How do we determine if either of these methods are effective? High and low levels of researcher-volunteer contact could both result in poor reporting. Therefore, a 'gold standard' method of reporting was implemented for comparison. This 'gold standard' consisted of people who were trained in game observing and were paid to go out into the community and report on games played by teams involved in the study. Table 2.1 summarizes the factors and their levels that were examined in this study.

Factor	Levels
Association Size	Large, Small
Contact	Gold Standard, High Contact, Low Contact
Age Division	Pee Wee, Bantam, Midget
Skill Level	C, B, A

Table 2.1: Factors and factor levels included in the study.

2.3 The ‘Gold Standard’

The ‘gold standard’, also referred to as the trained observers in this thesis, were students from Simon Fraser University who were hired to observe a sub-sample of games from teams participating in the study. Prior to observing a single game, they were trained, using video clips of minor hockey games, to detect an incident of concern. During each observed game, they recorded game information (team name, game date, etc.) as well as whether or not an incident occurred using the ‘observer game summary form’ (Figure 2.3). If no incident was observed the game was suspected to be incident free. For games in which an incident was observed, the observers were instructed to fill out the ‘Simon Fraser University Hockey Incident Recording Sheet’ (Figure 2.4). This sheet contained detailed information regarding the incident such as the time on the clock, score, area on the ice, player number, etc.

Ideally, the trained observers were to observe a simple random sample of games from participating teams in the study. This was accommodated as best as possible. In scheduling trained observers, their outside schedules and the fact that they could not always make it to assigned games was taken into account. Taking a simple random sample sometimes required observers to watch one game in one city and then travel to another city and watch another game. Sometimes, this method of sampling even required one observer to be in two places at once. Of course, this was not possible due to physical and temporal constraints. To solve this problem, ‘arena nights’ were created in which the observers were randomly assigned to a specific arena for a given night and would watch all the games of teams participating in the study. Reports were filed only for those teams that were involved in the study, resulting in several games in which only one team was observed. One reason for allowing the trained observers to only report incidents for participating teams was for the purpose of assessing volunteers. If 'gold standard' reports were made on teams with no volunteers, there would be no way to compare incidents reported by these two reporting methods within these games. If an observer was unable to make the assigned arena night, he/she was randomly re-assigned to another arena night that he could attend.

The factors age division, skill category and association size make up 18 (3x3x2) factor level combinations. A factorial design would require each combination to be observed at least once by the trained observers. Initially, it was planned to implement a time element such that trends in incident rate could be examined across regular season games, play-off races and play-offs. Therefore, it was desired to have each factor level combination observed once at each of these time points, which would require at least three observations in each of the 18 treatment combinations. However, scheduling complications, such as postponed games made observing each combination once per time frame difficult enough that it was not possible. This resulted in insufficient data to observe any trends in time. What was learned from this attempt to include time is that the only way to know with certainty when a team's next game will be is to be a part of that team as either a coach, player or parent. In fact, keeping abreast of the scheduling changes was so difficult, that, as the end of the season approached, we had yet to observe some level combinations. This created a dependency in the assignment of games as level combinations that needed to be sampled were sampled non-randomly.

2.4 Selecting Teams for Study Participation

Due to the size of BC, the logistics of this proof-of-concept study required that a region close to Simon Fraser University be be chosen from which to solicit volunteer participation. Consequently, only teams from the PCAHA were to be involved. As can be seen in Figure 2.2, this region encompassed five leagues in a heavily populated area. Two leagues were selected with the aim that every team within leagues would participate. This would result in a cluster sample where the clusters were the leagues.

By assuming that leagues within this region do not differ in their overall incident rate, cluster sampling would not affect any inferences we wished to make regarding the region as a whole. In terms of monetary and time cost, the size of this study was large enough within this region that any attempt to expand to other regions would have near impossible. The staff required to run this study within a single region consisted of three full time staff and trained observers who were paid $25 per game. To run this study in in another region would require three more full time staff and several more trained observers.

The two leagues selected to be clusters were the Lions Gate League and the Fraser Valley North League, containing a total of ten associations. Each league contained five associations and all were asked to participate in the study. One association chose not to participate, resulting in nine volunteer associations. Four associations were classified as large and five associations were classified as small. To balance the study design, two small associations (Whistler and Squamish) were combined to form a single, small association so that the final design had four large and four small associations.

Originally, it was assumed there was a possibility for interaction between volunteers within an association. To prevent this intereaction from affecting reporting rates, it was decided that each association, as a whole, be assigned one level of researcher-volunteer contact (high or low). By including four large and four small associations, we could randomly assign two associations from each group to a level of researcher-volunteer contact. This design balancing would prevent complications that may arise in the analysis had individual teams been randomly assigned to a level of contact.

A total of 150 teams were eligible to participate in this study. Of these, 90 teams volunteered to report incidents of concern. The process of recruiting teams to participate is a task that I do not wish assigned to my worst enemy simply because the process of contacting a representative from each team was so difficult. It is truly amazing how many busy people are involved in minor hockey.

In order to maintain a good standing relationship with all levels of the BCAHA, it was necessary to start at the top of the BCAHA political hierarchy and obtain contact information (phone number, e-mail and address) for each level below. This political hierarchy begins with the president of the BCAHA and branches out through lower executives ending in the individual team coaches and managers (Figure 2.5).

When contacting each level of the hierarchy, the study coordinator (Mr. Ian Williamson) would introduce himself and proceed to explain the objectives of the study. He would request support for the project in terms of permission to contact the next level of the hierarchy and would state that their contact information had been distributed with the support of the superior head in the association. He would also include that volunteer participation had been approved at all levels. The researchers felt that this process, although time consuming, resulted in a higher participation rate than if teams had simply been contacted without official approval from the association. This process was repeated until contact information for the coaching staff of each team was obtained.

In communicating with coaches or team managers, a brief 'sales pitch' was given along with a description of the type and amount of work that would be required from a volunteer. They had the opportunity to commit to the study or decline participation on behalf of the team. If willing, the coach would then designate a team volunteer and relay the contact information to Mr. Williamson who would then personally deliver a recording package to the recruited volunteer. This recording package was self-contained and informed the user how to record incidents of concern and how to submit the information to the researchers in the Sport-Concussion Research Group (SCRG) stationed out of the Motor Behaviour Laboratory. All volunteers were given a concussion recording form (Figure 2.6). These forms were used to record information about the volunteer (name, address, phone number, etc.), general information about the game (team, division, game date, etc.) and about the incident (cause, result, location on ice, etc.). The information was submitted via fax, e-mail, or mail and was required only for games in which an incident occurred and was submitted at the convenience of the volunteer. Once a form was submitted, a member of the SCRG would contact the volunteer and the submitted information was verified along with other information, such as who diagnosed the concussion.

The fact that only 60% of the teams contacted committed to the study indicates a possibility of self-selection bias. This would occur if the participants had a higher motivation to be in the study and if a sample of these motivated people would produce results that differed from the results obtained from a more general sample. In this study, it is possible that the coaches or managers who were willing to participate did so because they had more interest in concussions. People who are more aware or concerned about concussions may be more likely to report. This may result in a difference in reported incident rate than if all 150 teams had participated. Although the coaches or managers had the final say as to whether their team would participate or not, they were not the ones submitting reports to the SCRG. They were asked to recruit a parent volunteer to record and submit information. Therefore, the assumption could be made that the 60 teams not participating would not report any differently than the participating 90.

2.5 Data Collection

The number of incidents occuring in a game was recorded by volunteers for each team-game. A team-game is defined as a game played by a single team. One hockey game consists of two team-games. Each volunteer was therefore, only responsible for their own team. Likewise, trained observers were only required to report on teams participating in the study.

Initially, the study was designed using a two-phase sampling plan (see Lohr (1998) pg. 383) to gain a more precise estimate of the concussion rate. The first phase sample would consist of all games played by consenting teams. All of these games were to be monitored by parent volunteers. From this sample of games, trained observers would attend a sub-sample (phase II) of games. The incident rate as computed from the phase I sample would then be adjusted using the incident rates as reported by the volunteers and trained observers from the phase II sample. In the phase I sample, $\hat{t}_{vol}^{(1)}$ is the estimated incident total reported by the volunteers. From this sample, a sub-sample of games is taken in which trained observers attend and report incidents. From this phase II sample $\hat{t}_{obs}^{(2)}$ is the incident total reported by the trained observers and $\hat{t}_{vol}^{(2)}$ is the incident total reported by the volunteers. The estimate of the total number of incidents is found as

$$\hat{t} = \frac{\hat{t}_{vol}^{(1)}\hat{t}_{obs}^{(2)}}{\hat{t}_{vol}^{(2)}}$$

From this estimated total, the incident rate can be calculated by dividing by the total number of team-games observed. This proposed method of estimation is the rationale for the trained observers attending a simple random sample of games.

With consent from only 60% of the eligible teams, several factor level combinations were unable to be measured (Table 2.2). The trained observers were initially scheduled to visit all the observed factor combinations, but because of game changes and other uncontrollable forces, only 80 of the 90 consenting teams were observed by the trained observers. Between 1 and 5 games of each team were visited by trained observers (Table 2.3)

Association Size	Contact	Division	Skill Level
Large	High	Bantam	B
Large	High	Midget	B
Small	High	Pee Wee	B
Small	High	Bantam	B
Small	Low	Pee Wee	B

Table 2.2: Missing factor level combinations due to lack of volunteer participation.

Number of Games for a Particular Team	Number of Teams	Total Team Games Monitored
1	37	37
2	27	54
3	12	36
4	2	8
5	2	10
Total	80	145

Table 2.3: Frequencies of number of observations on a given team by the trained observers.

It was our aim that in sending out the trained observers to obtain a 'gold standard' estimate of the concussion rate, the majority of their reported incidents would have coincided with the reports filed by the volunteers, who were assumed to be at those same games. This would have provided us with some sense of the performance of the volunteers with respect to reporting incidents. This, in fact, did not happen. Of the 8 incidents reported by the trained observers, not one was also reported by the volunteers. Of the 31 incidents reported by the volunteers, none of these were captured in the sub-sample collected by the trained observers. The incident reports filed by the volunteers highlighted the fact that the number of trained observers available were more limited than anticipated and in fact may not have been the appropriate 'gold standard' for comparison.

Observer Game Summary Form

Observer Information

Name: ______________________ Phone / Email: ________________________

Game
Date: Day: __________ Month: ___________ Year: __________

Home Team: ______________________ Visiting Team: ____________________________

League Information

Division of Play. (Please check one)

☐ Midget ☐ Bantam ☐ Pee Wee

Category of Play. (Please check one)

☐ AAA ☐ AA ☐ A ☐ BB ☐ B ☐ C

City / Home Rink: ____________________________________

Was there an incident of concern during this game?

☐ Yes ☐ No If yes, please complete the remainder of this form and the Incident Recording Sheet.

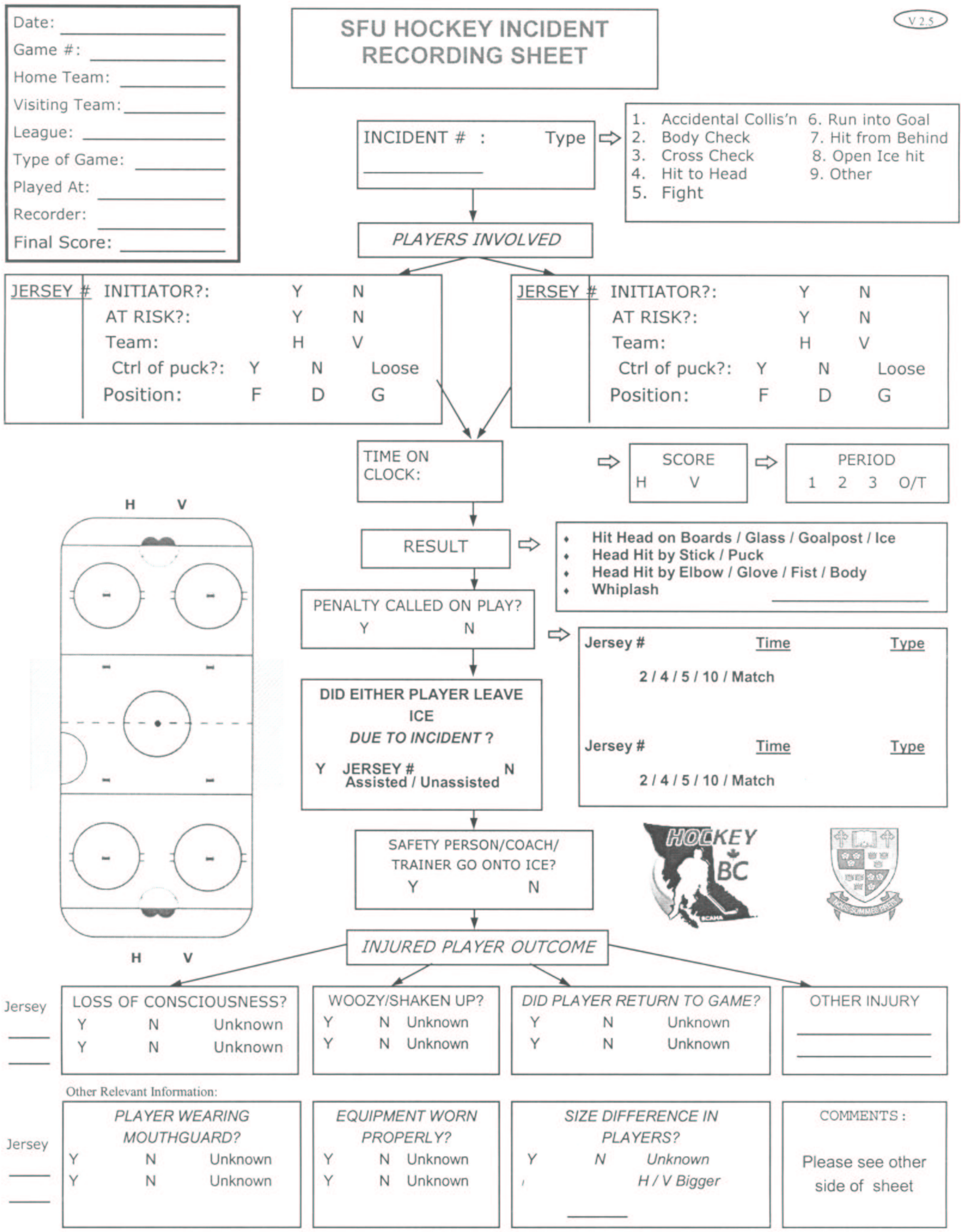
Incident Description

Please briefly note the aspects of the incident that were suggestive of a concussion.
* What compelled you to file an Incident Recording Sheet? *

Things to consider include, but are not exclusive to:
Loss of consciousness; Player did not immediately resume game play; Player attended to on ice / bench; Player woozy / shaken up; Player did not return to game; Unfavorable change in game play.

Remarks

Figure 2.3: Summary form submitted by trained observers after each game.



SFU HOCKEY INCIDENT RECORDING SHEET

V 2.5

Date: ______
Game #: ______
Home Team: ______
Visiting Team: ______
League: ______
Type of Game: ______
Played At: ______
Recorder: ______
Final Score: ______

INCIDENT # : Type ______

1. Accidental Collis'n
2. Body Check
3. Cross Check
4. Hit to Head
5. Fight
6. Run into Goal
7. Hit from Behind
8. Open Ice hit
9. Other

PLAYERS INVOLVED

JERSEY #				
	INITIATOR?:	Y	N	
	AT RISK?:	Y	N	
	Team:	H	V	
	Ctrl of puck?:	Y	N	Loose
	Position:	F	D	G

JERSEY #				
	INITIATOR?:	Y	N	
	AT RISK?:	Y	N	
	Team:	H	V	
	Ctrl of puck?:	Y	N	Loose
	Position:	F	D	G

TIME ON CLOCK:

SCORE
H V

PERIOD
1 2 3 O/T

H V

H V

RESULT

- Hit Head on Boards / Glass / Goalpost / Ice
- Head Hit by Stick / Puck
- Head Hit by Elbow / Glove / Fist / Body
- Whiplash ______

PENALTY CALLED ON PLAY?
Y N

Jersey #	Time	Type
2 / 4 / 5 / 10 / Match		
Jersey #	Time	Type
2 / 4 / 5 / 10 / Match		

DID EITHER PLAYER LEAVE ICE *DUE TO INCIDENT?*
Y JERSEY # N
Assisted / Unassisted

SAFETY PERSON/COACH/TRAINER GO ONTO ICE?
Y N

HOCKEY BC

INJURED PLAYER OUTCOME

Jersey	LOSS OF CONSCIOUSNESS?	WOOZY/SHAKEN UP?	*DID PLAYER RETURN TO GAME?*	OTHER INJURY
____	Y N Unknown	Y N Unknown	Y N Unknown	______
____	Y N Unknown	Y N Unknown	Y N Unknown	______

Other Relevant Information:

Jersey	*PLAYER WEARING MOUTHGUARD?*	*EQUIPMENT WORN PROPERLY?*	*SIZE DIFFERENCE IN PLAYERS?*	COMMENTS:
____	Y N Unknown	Y N Unknown	*Y N Unknown*	Please see other side of sheet
____	Y N Unknown	Y N Unknown	*H / V Bigger* ____	

Figure 2.4: Incident information recorded by trained observers.

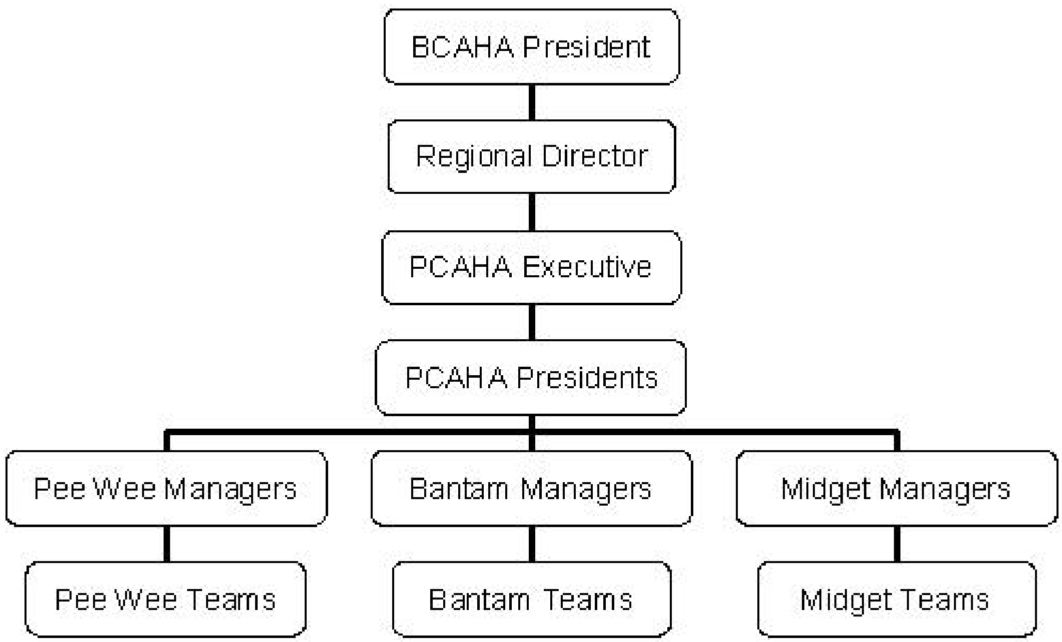


Figure 2.5: Minor hockey political hierarchy. Each level in the hierarchy had to be contacted prior to contacting the subsequent level.

Concussion Recording Form

Recorder Information

Name: ____________________ Phone / Email: ____________________

What team are you associated with?

____

Reporting Date: Day: ________ Month: ________ Year: ________

League Information

Division of Play: (Please check one)

☐ Midget ☐ Bantam ☐ Pee Wee

Category of Play: (Please check one)

☐ A ☐ B ☐ C

City / Home Rink: (Please check one)

☐ Hollyburn ☐ North Vancouver ☐ West Vancouver ☐ Squamish ☐ Whistler

☐ Coquitlam ☐ Port Coquitlam ☐ Port Moody ☐ Ridge Meadows

Game Information

Date of Injury: Day: ________ Month: ________ Year: ________

Home Team: __________ Visiting Team: __________

Were you in attendance?

☐ Yes ☐ No

Injury Data

Cause of Incident: (Please check one)

☐ Accidental Collision ☐ Body Check ☐ Cross Check ☐ Hit to Head

☐ Run into Goal ☐ Hit from Behind ☐ Open Ice Hit ☐ Other

Result of Incident: (Please check one)

Hit Head on: or, Hit Caused:

☐ Boards ☐ Glass ☐ Goalpost ☐ Ice ☐ Whiplash ☐ Other

or, Head Hit by:

☐ Elbow ☐ Glove ☐ Fist ☐ Body ☐ Stick ☐ Puck

Location of Incident With Respect to Injured Player: (Please check one)

☐ Offensive Left Corner ☐ Offensive Right Corner ☐ Offensive Slot ☐ Neutral Zone

☐ Defensive Left Corner ☐ Defensive Right Corner ☐ Defensive Slot

Remarks

Figure 2.6: Information sheet submitted by the volunteers when an incident of concern was observed.

Chapter 3

The Model

3.1 Overview of Analysis

The data was analyzed by applying generalized estimating equations (GEE's) to a candidate set of log-linear models and then, to account for model-to-model variation, averaging the model estimates using the Akaike Information Criterion (AIC). At first glance it may appear that these two methods conflict since GEE's are not based on maximum likelihood techniques but are more closely related to the quasi-likelihood approach (Wedderburn, 1974, McCullagh,1983) in which a distribution need not be specified. On the other hand, model averaging using AIC is based on maximum likelihood estimation, which requires a specified distribution. However, an adjustment to the AIC value can be made using the estimated scale parameter to keep these two techniques from conflicting. This adjustment allows for model averaging to be used in conjunction with GEE's.

As is common with count data, it is assumed that the number of incidents in a game follows a Poisson-like distribution with rate parameter λ and a possible scale parameter ϕ to describe any over-dispersion. Observations were taken on individual team-games. Each game was assumed to consist of two independent team games. This assumption was necessary due to games where only one of the teams was a participant in the study. Also, because each participating team had their own volunteer data recorder, incidents sustained by one team were assumed to have no effect on the reports filed by the volunteers from the opposing team since, to facilitate reporting, each team was required only to report on their own players. However, there were a number of team games in which reports were filed by both a volunteer observer and a trained observer. Having two reports for the same game suggested that there should be a dependence between them.

Considering each team game as a cluster with a maximum of two observations, the joint distribution of observations within a cluster is unknown. Therefore, a potential analysis for this data is one in which a distribution need not be specified. The GEE's proposed by Liang and Zeger (1986) allow for distributional assumptions to be made on the marginal data (reports by individuals) but only weak assumptions regarding the mean-variance relationship of the joint distribution (reports from the same team game). Applying this method to a log-linear model will produce estimates for the rate of concussions/team-game for each factor level combination.

3.2 General Estimating Equations

Let the observations of the number of incidents on a team game be noted by the vector $\mathbf{Y}_i = (y_{i1}, \ldots, y_{in_i})^T$ where for the i^{th} subject or cluster $(i = 1, \ldots, K)$, there are a total of n_i observations $(n_i \leq 2 \ \forall \ i)$. Associated with each observation is a vector of covariates $\mathbf{X}_{it} = (x_{it1}, \ldots, x_{itp})^T$ representing the factors describing the team game. These covariates or factors are represented as an $n_i \times p$ matrix for each of the K clusters $\mathbf{X}_i = (x_{i1}, \ldots, x_{in_i})^T$. Since count data often displays greater variability than the Poisson assumption allows, the marginal distribution can be adjusted using the scale parameter ϕ (Agresti, 2002). The marginal distribution for each $y - it$ is

$$f(y_{it}) = exp[\{y_{it}\theta_{it} - a(\theta_{it}) + b(y_{it})\}\phi] \tag{3.1}$$

Let $\theta_{it} = h(\eta_{it})$ and $\eta_{it} = \mathbf{X}_{it}\beta$ where $\mathbf{h}(\cdot)$ is termed the link function such that it 'links' the parameter of the distribution to several predictor variables by way of a vector of parameters β. It is straight forward to show that using this exponential family form, the expectation and variance of y_{it} are

$$E(Y_{it}) = a'(\theta_{it}) \qquad \text{and} \qquad V(Y_{it}) = a''(\theta_{it})/\phi.$$

When estimating the parameters β, using maximum likelihood estimation, it is common to set the derivative of the natural logarithm of the joint distribution with respect to each β_j to 0 for some $\hat{\beta}_j$, thereby maximizing β_j with respect to the data at hand. Note that setting this derivative equal to 0 may only find a local maximum or even a minimum and therefore, one needs to ensure that $\hat{\beta}_j$ is indeed a maximum. However, in the case of the exponential family, $\hat{\beta}_j$ is the guaranteed maximum. Using the chain rule, the maximum likelihood estimating equations for each β_j are:

$$\frac{\partial l(\theta_{it})}{\partial \beta_j} = \frac{\partial l(\theta_{it})}{\partial \theta_{it}} \frac{\partial \theta_{it}}{\partial \eta_{it}} \frac{\partial \eta_{it}}{\partial \beta_j} = 0.$$

In matrix notation, considering the entire vector of parameters β,the estimating equations can be expressed as

$$U(\beta) = \sum_{i=1}^{K} X_i^T \Delta_i S_i = 0 \tag{3.2}$$

where $\Delta_i = diag(\partial\theta_{it}/\partial\eta_{it})$ and $S_i = Y_i - a'(\theta_i)$. If A_i is defined as the diagonal matrix $diag(a''(\theta_{it}))$, the solution to these equations $\hat{\beta}$ has been shown to be consistent estimators of β by Liang and Zeger (1986) with asymptotic variance

$$V = \left(\sum_{i=1}^{K} X_i^T \Delta_i A_i \Delta_i X_i \right)^{-1} \left(\sum_{i=1}^{K} X_i^T \Delta_i Cov(Y_i) \Delta_i X_i \right) \left(\sum_{i=1}^{K} X_i^T \Delta_i A_i \Delta_i X_i \right)^{-1}$$

This variance, termed the 'sandwich estimator' corrects for model mispecification by sandwiching the variance of β calculated from the data between two estimates of the information matrix, calculated under certain model assumptions (Hardin and Hilbe, 2003).

However, observations made on the same subject or cluster are generally correlated with each other. Treating correlated observations as independent decreases the efficiency of the resulting estimates (Liang and Zeger, 1986). Liang and Zeger (1986) also proposed that, using GEE's, one can incorporate the correlation structure of the data into the estimating equations 3.2. To accomplish this, let $V_i = A_i^{1/2} R(\alpha) A_i^{1/2}/\phi$ where $R(\alpha)$ is a correlation matrix fully specified by the vector of parameters α. Note that if $R(\alpha)$ is the true correlation structure then $V_i = Cov(Y_i)$. The parameters β can then be estimated after the method of moments estimates of α and ϕ are substituted into $R(\alpha)$ and ϕ respectively. This results in the estimating equations

$$U(\beta) = \sum_{i=1}^{K} D_i^T V_i^{-1} S_i = 0 \tag{3.3}$$

where $D_i = A_i \Delta_i X_i$ Therefore,

$$U(\beta) = \sum_{i=1}^{K} U_i \left[\beta, \hat{\alpha}\{\beta, \hat{\phi}(\beta)\right] = 0$$

The solution to these equations $\hat{\beta}$ have also been shown to be consistent (Liang and Zeger, 1986) with asymptotic variance

$$V = \left(\sum_{i=1}^{K} D_i^T V_i^{-1} D_i\right)^{-1} \left(\sum_{i=1}^{K} D_i^T V_i^{-1} Cov(Y_i) V_i^{-1} D_i\right) \left(\sum_{i=1}^{K} D_i^T V_i^{-1} D_i\right)^{-1}$$

This 'sandwich estimator' is modified such that it incorporates a specified working correlation structure that the data is assumed to follow. The middle term uses the empirical correlation to adjust the estimate for a mis-specified working correlation. This estimated variance is consistent for estimates of β (Liang and Zeger, 1986). Furthermore, simulation studies by Liang and Zeger (1986) demonstrated that, although comparable for low correlated data, as the correlation between observations increases (α increased from 0.3 to 0.7), the GEE estimates were always more efficient relative to the independence estimating equations if the correct correlation structure was specified. Efficiency of the GEE estimates increases if the working correlation structure is close to the true correlation structure. It was also demonstrated that the relative efficiency of the GEE estimator also increases if the number of observations for each cluster varies.

For independent observations, $\hat{\beta}$ is estimated using the Gauss-Newton iterative algorithm. The same algorithm is used for dependent observations except that the algorithm incorporates estimates of α and ϕ. The algorithm is as follows

$$\begin{aligned}\hat{\beta}_m &= \hat{\beta}_{m-1} - \left(\sum_{i=1}^{K} D_i^T(\hat{\beta}_{m-1})\tilde{V}_i^{-1}(\hat{\beta}_{m-1})D_i^T(\hat{\beta}_{m-1})\right)^{-1} \\ &\times \left(\sum_{i=1}^{K} D_i^T(\hat{\beta}_{m-1})\tilde{V}_i^{-1}(\hat{\beta}_{m-1})S_i(\hat{\beta}_{m-1})\right)\end{aligned}$$

where $\tilde{V}_i(\beta) = V_i(\beta, \hat{\alpha}(\beta, \hat{\phi}(\beta)))$ (i.e. a function of β after α and ϕ have been estimated). Note that these results are generalized, not only for any member of the exponential family but for any distribution in which a mean-variance relation is specified.

There are several options available for estimating the correlation between longitudinal observations. The exchangeable correlation structure specifies a constant correlation α for all pairs of observations. The autoregressive structure specifies the correlation as a decreasing function of the time interval between any two observations. An unstructured correlation specifies a different correlation parameter α_{ij} for all possible i, j pairs of observations. Many other structures exist and may be preferable depending on the structure of the data. For collected data in which there is a maximum of two observations per cluster, all structures will provide the same results.

3.3 Akaike Information Criterion (AIC)

Model selection methods are commonly based on hypothesis tests using a ratio of likelihood functions. The hypothesis tests generally consist of comparing two models; one with fewer parameters than the other. The likelihood ratio statistic is asymptotically distributed as a χ^2 random variable with degrees of freedom equal to the difference in number of parameters between the two models (Agresti, 2002). Hypothesis testing is dependent on the subjective Type I error level imposed on each test. Problems arise when many tests are performed as this acts to inflate the experimentwise Type

I error. Further problems with this method are due to the χ^2 approximation of the likelihood ratio. If models are not nested, they may have equal number of parameters such that their difference is zero (a χ^2 random variable can not have 0 degrees of freedom). Thus, the test statistic's distribution is unknown and more difficult to find.

Akaike (1978) introduced a method of model selection, the AIC, which makes use of the Kullback-Leibler information. This process allows for non-nested models to be compared and does not inflate Type I error resulting in the ability to examine many different models at once. The K-L information is a measure of the distance between two functions. For our purposes, the functions are labeled $f(x)$, the true function, and $g(x)$, an estimate of $f(x)$. The K-L information is given by

$$I(f(x), g(x)) = \int f(x) log(\frac{f(x)}{g(x)})\, dx \tag{3.4}$$

The function that is closest to the true function $f(x)$ will minimize the K-L information. In terms of model selection, let both $f(x)$ and $g(x)$ be probability distributions. As an estimated distribution, $g(x)$ is known to rely on a set of parameters given by θ_0. The true distribution, $f(x)$, is unknown preventing the calculation of the exact K-L information. However, by rewriting 3.4 as

$$I(f(x), g(\theta_0|x)) = \int [log(f(x))] f(x)\, dx - \int log(g(\theta_0|x)) f(x)\, dx$$

Then,

$$I(f(x), g(\theta_0|x)) = E_f[log(f(x))] - E_f[log(g(\theta_0|x))] \tag{3.5}$$

we can see that the first term in 3.5 is a constant and in order to minimize 3.5 we must maximize the second term. The true value of the probability function $g(\theta_0|x)$ will minimize the K-L information for all $\theta \in \Theta$. In practice, the true value of a distribution function is never known and therefore must be estimated. The maximum likelihood estimate $\hat{\theta}$ can be used to estimate $g(\theta_0|x)$. By using $\hat{\theta}$ to minimize the K-L information we are minimizing its expected value instead of its true value. Akaike (1973) showed that an unbiased estimate of the second term in 3.5 is $log(g(\hat{\theta}|x)) - K$

where K is the number of parameters estimated in $g(\hat{\theta}|x)$. For historical purposes, this estimate was multiplied by -2 so that

$$\text{AIC} = -2log(g(\hat{\theta}|x)) + 2K$$

Therefore, AIC allows for potential models to be selected based on maximum likelihood methods by choosing the model with the minimum AIC value. It also incorporates the principle of parsimony in that the likelihoods are penalized for having too many parameters, thereby limiting their final number.

Using the AIC for model selection requires maximum likelihood techniques and, therefore, certain model assumptions must be made regarding the distribution of the data. Generalized estimating equations do not require that any distribution be specified; only a mean and variance relationship need be assumed. To address the use of AIC with respect to generalized estimating equations and other quasi-likelihood techniques, Burnham and Anderson (1998) suggested the use of an adjusted AIC which incorporates the estimated scale parameter ϕ. Variance estimates in this project are multiplied by ϕ for over-dispersion adjustment. The adjusted AIC in this situation is termed QAIC (for quasi-likelihood) and the QAIC value to be used for model selection is

$$-2log(g(\hat{\theta}|x))\phi + 2K$$

where K is the number of estimated parameters plus one for the scale parameter. Burnham and Anderson (1998) also suggest that when comparing candidate models, they all be fitted with the same scale parameter that was estimated for the global or fullest model. This will ensure that all variance estimates are adjusted equally and will reduce model to model variation. Both AIC and QAIC act to penalize good models that are over-parameterized in order to restrict the number of these types of models in the competing set.

There may be cases when several models result in similar AIC values. Since the K-L information can only be estimated up to the value of an unknown constant, only relative AIC values between competing models can be examined. Given two models

$g_1(x)$ and $g_2(x)$, we want to select the model such that $I(f(x), g_1(x)) < I(f(x), g_2(x))$. But how close is $g_1(x)$ to $f(x)$? Both of these models could be poor estimates of $f(x)$, or both could be good estimators of $f(x)$. Using the notion of QAIC differences (ΔQAIC), one can order the set of candidate models relative to each other and examine how well the models perform relative to each other and relative to the best model in the candidate set. Models with large values of ΔQAIC are considered to have little empirical support. However, if there are several ΔQAIC values that are close to 0 (where 0 implies the best candidate model in the set), how does one decide which model is the best? This leads to the notion of model averaging in which estimates from several valid candidate models are averaged to account for sample-to-sample model variation. For example, given a sample y, $g_1(y)$ might be better (lower ΔQAIC) than $g_2(y)$. However, in another independent sample from the same population, $g_2(x)$ might be a better fit than $g_1(x)$. Using the notion of ΔQAIC for a given candidate model, one can compute the Akaike weight for that model as a measure of the evidence that the model is the best model for the given data. Akaike weights are calculated using the equation

$$w_i = \frac{exp(-\frac{1}{2}\Delta_i)}{\sum_{i=1}^{R} exp(-\frac{1}{2}\Delta_i)} \tag{3.6}$$

where Δ_i is the difference between the ΔQAIC for model i and the minimum QAIC value in the set of models. The numerator in 3.6 is proportional to the ratio of the adjusted likelihoods for model i and the model with the lowest QAIC value:

$$\begin{aligned} exp(-\frac{1}{2}\Delta_i) &= exp(-\frac{1}{2}[QAIC_i - QAIC_{min}]) \\ &= exp(-\frac{1}{2}[2Log(L_i) - 2K_i - (2Log(L_{min}) - 2K_{min})]) \\ &= exp(-\frac{1}{2}[2Log(L_i) - 2Log(L_{min}) - (2K_i + 2K_{min})]) \\ &\propto exp(-\frac{1}{2}[2Log(L_i) - 2Log(L_{min})]) \end{aligned} \tag{3.7}$$

These values are than normalized such that all weights in the current set of models sums to 1. It is clear that as Δ_i increases, w_i decreases providing less evidence that model i is best. If models are dropped or added to the set of candidate models, both $\Delta_i's$ and $w_i's$ need to be recalculated.

Burnham and Anderson (1998) discuss that in order to account for variation in estimates among competing models, one should form an average of estimates over the set of all R candidate models. The proposed method for averaging parameters makes use of the i^{th} model's Akaike weight and whether or not the j^{th} parameter is in that model by incorporating an indicator function $I_{[\beta_j \epsilon g_i(\cdot)]}$ that equals 1 if β_j is in model i and 0 otherwise. Then

$$\hat{\hat{\beta}}_j = \sum_{i=1}^{R} w_i \hat{\beta}_{i,j} I_{[\beta_j \epsilon g_i(\cdot)]}$$

is an average of the $\beta_j's$ adjusted by the total weights of all models that incorporate variable x_j. The variance of these averaged parameters and resulting fitted response values is given by:

$$Var(\hat{\hat{\theta}}) = \left[\sum_{i=1}^{R} w_i \sqrt{var(\hat{\theta}_i|g_i) + (\hat{\theta}_i - \hat{\hat{\theta}})} \right]^2$$

This variance incorporates the standard errors of the estimates conditional on the model and averages the R standard errors according to their associated weights.

3.4 Model Fitting

Data analysis was performed using the SAS GENMOD procedure (SAS/STAT software,Version 8, Copyright ©1999, SAS Institute INC). This procedure allows for the GEE analysis proposed by Liang and Zeger (1986). Standard errors were calculated using the empirical values which do not rely on distributional assumptions. The candidate model set consisted of 16 models, and a weighted average of their parameters was computed based on their associated Akaike weights.

Chapter 4

Results

4.1 Summary Statistics

By the end of the 2002-2003 hockey season, almost 40 'head-incidents of concern' were reported from over 1900 team-game observations. Trained observers attended 143 team-games, high contact volunteers observed 865 team-games and low contact volunteers observed 908 team-games for a total of 1773 independent team-games and 1916 observations. There were 39 total reports between the three levels of contact. Trained observers reported 8 incidents of concern in contrast with the high contact volunteers who reported 22 incidents and low contact volunteers who reported 9 incidents. In terms of age division, the majority of incidents were reported in the pee wee division (17) with 9 and 13 reported in bantam and midget divisions respectively. The lowest skill level, C, had the highest number of incidents (28) which was almost 5 times higher than the highest skill level (6) and the B level teams (5). When association size was examined, the number of incidents reported by large associations was almost three times higher then in small associations (29 amd 10, respectively)

Empirical rates indicate that trained observers reported the highest incident rate for almost all of the main effects (Table 4.1). The one exception noted was the rate for large associations. Here, the trained observers reported 0.05 incidents/team-game whereas the high contact volunteers reported 0.062 incidents/team-game.

Volunteers reports proved to be more definitive in concussion confirmation as Table

4.2 highlights that 52% of incidents were diagnosed by a physician. Seven incidents reported by high contact volunteers and nine reported by low contact volunteers were physician confirmed concussions. Trained observers had no follow-up ability.

4.2 Model and Model Estimates

The over-dispersion parameter for the global model was estimated to be 1.21 indicating that the global model was an adequate fit to the data. Four models from the candidate set have a ΔQAIC value less than three (Table 4.3) suggesting substantial empirical support for these models (Burnham and Anderson, 1998). Eight models have a ΔQAIC greater than 10 suggesting that there is almost no empirical support for these models. Table 4.4 lists the averaged parameter estimates for all models in the candidate set, their Akaike weight, associated Z-scores and p-values. Each parameter estimate represents the natural logarithm of the ratio of the rates between two levels of a factor when all other factor levels are held the same. For example, the parameter estimate of 1.72 for the comparison of trained observers to low contact volunteers indicates that, on average, trained observers report $\epsilon^{1.72} = 5.58$ more incidents than the low contact volunteers. In terms of level of researcher-volunteer contact, both trained observers and high contact level volunteers had reporting rates that were significantly different than the low contact level volunteers ($p = 0.0005$ and $p = 0.02$ respectively).

Tables 4.5 to 4.7 list the fitted incident rates for each factor level combination that data was collected on. Figure 4.1 displays these rates graphically along with a common 95% confidence interval. The points to the far right of the graph represent the 'gold standard' estimates and the points to the far left represent the low contact level volunteer estimates.

In determining which contact method was the most optimal relative to the trained observers, a post-hoc comparison of the trained observers with the high contact level suggested that they were significantly different in their reporting rates (p=0.03). On average, the trained observers reported a rate 2.2 times that of the high contact level volunteers.

Main Effect	Level	Games Observed	Incidents of Concern	Empirical Estimate Incidents/Team-Game
		Low Contact Level		
	Pee Wee	257	2	0.008
Division	Bantam	361	0	0.000
	Midget	290	7	0.024
	C	594	1	0.002
Level	B	163	3	0.018
	A	151	5	0.033
Size	Small	357	2	0.006
	Large	551	7	0.013
		High Contact Level		
	Pee Wee	414	13	0.031
Division	Bantam	260	4	0.015
	Midget	191	5	0.026
	C	673	22	0.033
Level	B	63	0	0.000
	A	129	0	0.000
Size	Small	589	5	0.008
	Large	276	17	0.062
		Trained Observers		
	Pee Wee	60	2	0.033
Division	Bantam	51	5	0.098
	Midget	32	1	0.031
	C	106	5	0.047
Level	B	16	2	0.125
	A	21	1	0.048
Size	Small	43	3	0.070
	Large	100	5	0.050

Table 4.1: Number of games observed and empirical incident rates for each level of researcher-volunteer contact for each main effect.

Diagnosed By	Number	Percent Total
Doctor	16	52.0
Coach	6	19.0
Safety Person	2	6.5
Other	2	6.5
Unspecified	5	16.0
Total	31	100.0

Table 4.2: Breakdown of incidents of concern reported and to what level they were examined

	Factors In the Model						
Model	Size	Researcher-Volunteer Contact	Age-Division	Skill Level	Number of Parameters	ΔQAIC	Akaike Weights
1		x			4	0.00	0.309
2	x	x			5	0.68	0.234
3		x	x		6	1.14	0.201
4	x	x	x		7	2.25	0.120
5		x		x	6	3.78	0.047
6	x	x		x	7	4.46	0.036
7		x	x	x	8	4.92	0.031
8	x	x	x	x	9	6.04	0.018
9		Intercept Only			2	10.31	0.001
10	x				3	10.33	0.001
11			x		4	11.23	0.000
12	x		x		5	11.91	0.000
13				x	4	14.31	0.000
14	x			x	5	14.33	0.000
15			x	x	6	15.23	0.000
16	x		x	x	7	15.68	0.000

Table 4.3: Set of competing models with ΔQAIC and Akaike weights

Parameter	Estimate	Standard Error	Z-value	p-value
Large vs Small Association	0.14	0.16	0.91	0.18
Trained Observers vs Low Contact Level	1.72	0.52	3.27	0.0005
High Contact vs Low Contact Level	0.93	0.46	2.01	0.02
Pee Wee vs Midget	-0.09	0.14	0.67	0.25
Bantam vs Midget	-0.23	0.15	1.51	0.07
C vs A	-0.01	0.07	0.13	0.45
B vs A	0.02	0.095	0.18	0.43

Table 4.4: Akaike model averaged estimates and standard errors. Note that the intercept represents the baseline of small sized, low contact level, midget A teams.

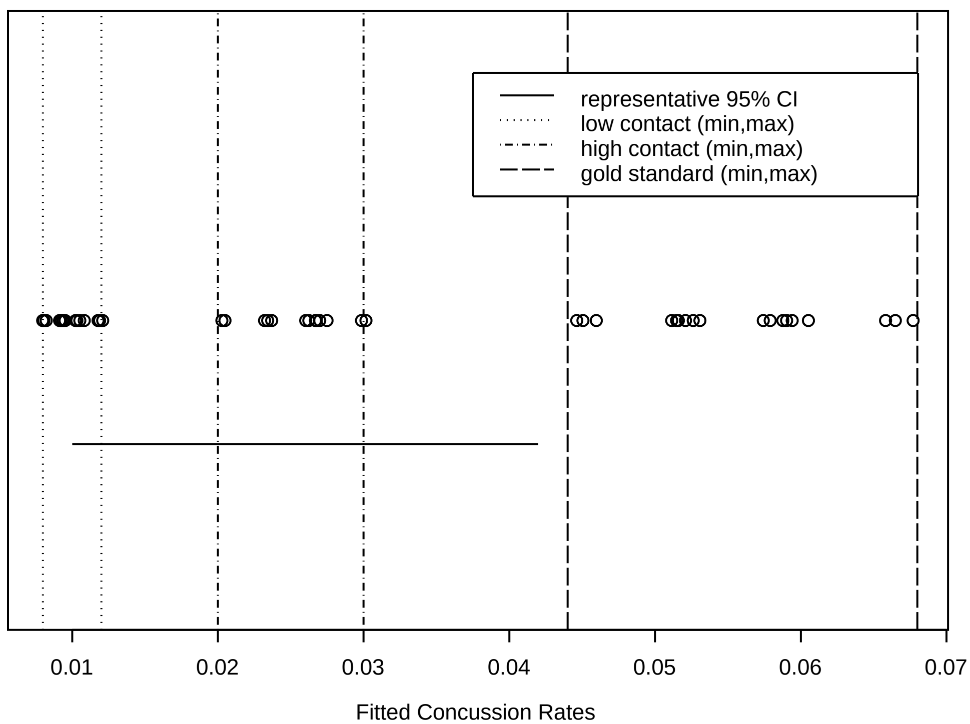


Figure 4.1: Graphical representation of the estimated rates for each factor level combination

Size	Contact Level	Division	Estimated Incidents/ Team-Game	95% CI
		Pee Wee	0.009	(0.004,0.021)
	Low	Bantam	0.008	(0.004,0.017)
		Midget	0.010	(0.004,0.026)
		Pee Wee	0.023	(0.012,0.045)
Small	High	Bantam	0.020	(0.011,0.039)
		Midget	0.026	(0.014,0.049)
		Pee Wee	0.051	(0.023,0.114)
	Gold Standard	Bantam	0.044	(0.019,0.111)
		Midget	0.057	(0.026,0.131)
		Pee Wee	0.011	(0.004,0.024)
	Low	Bantam	0.009	(0.005,0.020)
		Midget	0.012	(0.005,0.029)
		Pee Wee	0.027	(0.014,0.050)
Large	High	Bantam	0.023	(0.013,0.043)
		Midget	0.030	(0.017,0.053)
		Pee Wee	0.059	(0.029,0.118)
	Gold Standard	Bantam	0.051	(0.024,0.113)
		Midget	0.066	(0.032,0.135)

Table 4.5: Model-averaged concussion rates/team-game and 95% confidence intervals for skill level C

Size	Contact Level	Division	Estimated Incidents/ Team-Game	95% CI
Small	Low	Pee Wee	0.009	(0.004,0.022)
		Bantam	0.008	(0.004,0.019)
		Midget	0.011	(0.004,0.027)
	High	Pee Wee	-	(-,-)*
		Bantam	-	(-,-)*
		Midget	0.027	(0.014,0.052)
	Gold Standard	Pee Wee	0.053	(0.023,0.125)
		Bantam	0.046	(0.018,0.122)
		Midget	0.059	(0.025,0.144)
Large	Low	Pee Wee	0.011	(0.004,0.027)
		Bantam	0.010	(0.004,0.023)
		Midget	0.012	(0.005,0.032)
	High	Pee Wee	0.027	(0.014,0.054)
		Bantam	-	(-,-)*
		Midget	-	(-,-)*
	Gold Standard	Pee Wee	0.061	(0.028,0.132)
		Bantam	0.053	(0.023,0.128)
		Midget	0.068	(0.031,0.150)

Table 4.6: Model-averaged concussion rates/team-game and 95% confidence intervals for skill level B. *missing values indicate missing factor combinations

Size	Contact Level	Division	Estimated Incidents/ Team-Game	95% CI
		Pee Wee	0.009	(0.004,0.020)
	Low	Bantam	0.008	(0.003,0.020)
		Midget	0.010	(0.004,0.029)
		Pee Wee	0.023	(0.012,0.046)
Small	High	Bantam	0.020	(0.011,0.040)
		Midget	0.026	(0.014,0.051)
		Pee Wee	0.052	(0.023,0.120)
	Gold Standard	Bantam	0.045	(0.018,0.115)
		Midget	0.058	(0.025,0.140)
		Pee Wee	-	(-,-)*
	Low	Bantam	0.009	(0.004,0.025)
		Midget	0.012	(0.004,0.035)
		Pee Wee	0.027	(0.014,0.054)
Large	High	Bantam	0.024	(0.013,0.046)
		Midget	0.030	(0.016,0.059)
		Pee Wee	0.059	(0.028,0.130)
	Gold Standard	Bantam	0.052	(0.023,0.124)
		Midget	0.066	(0.031,0.150)

Table 4.7: Model-averaged concussion rates/team-game and 95% confidence intervals for skill level A. *missing values indicate missing factor combinations.

Chapter 5

Recommendations

In retrospect, this study had three major flaws.

1. The term incident of concern was too subjective.
2. The 'gold standard' was not a true 'gold standard'.
3. The current study design does not allow for cost effective expansion to other regions of BC.

5.1 Measurement Subjectivity

'Incident of concern' is a very subjective term and, even if it is clearly defined, has the potential for different people to provide different reports. One person's idea of a potential concussion may be different from another's resulting in two people perceiving the same event differently. This can lead to high variability between observers (both trained and volunteer). It is desirable to have a response variable that is clear to all observers such that this variability is minimized.

It is recommended that future studies focus primarily on concussions. Previously, reports were filed if there was a concern regarding a potential concussion but there was no real evidence to back up those concerns other than the incident itself. Although reports received from people who had actually taken players to a physician was beneficial, this may not happen as often as we would like. Therefore, if the volunteer is provided with a symptom checklist, they can make a more informed decision in terms of filing a report or not. If an incident occurs and the affected player exhibits one or more of the listed symptoms, than the volunteer is asked to file a report. By providing a list, our aim is that if the volunteer is concerned enough to file a report based on the present symptoms, they will ensure that the player visits a physician for confirmation. This will provide us with the confidence to say that these reports were indeed concussions.

5.2 Gold Standard Limitations

Keeping the trained observers independent from the volunteers made post-incident player follow-up impossible with respect to the trained observers. This resulted in a lowered ability to compare volunteer reports with the 'gold standard'. Since none of the volunteer reports matched any of the eight, trained observer reports, we can assume that there are three possibilities for the team-games attended by the trained observers. The first possibility is that no incident occured. If there was no volunteer report and no trained observer report than we were more confident in our assumption that nothing happened. The second possibility is an incident occurred and it was reported by the trained observer but not the volunteer. If there was no subjectivity in the term incident of concern than this possibility would help us meet Objective 2 of this study. The third possibility is that an incident was reported by a trained observer because they may have seen a player get hit, which resulted in his head hitting the side boards before falling to the ice. From the stands, this would appear to be an incident but perhaps the player only hurt their knee or another part of their body not associated with their head. Only post-incident follow-up would be able to determine if this was indeed a true incident.

To truly have a 'gold standard' with which to compare volunteer reports, interaction with teams needs to be allowed. If a trained observer can attend team-games in which they are able to communicate with the parents or coaches afterwards to determine if any concussions may have occurred. If possible, a trained observer should follow a team throughout the entire season. I feel that this constant communication and resulting familiarity will increase the diligence of these teams to file concussion reports. Sampling team-games will no longer be a simple random sample but a cluster sample where every team-game from a given team (cluster) is observed. If the assumption is made that teams of the same age division and skill level from different associations have similar concussion rates, then using a cluster sample should not effect the precision of the estimates.

5.3 Cost Effective Study Expansion

Since the 'gold standard' is currently unable to determine whether parent volunteers are an adequate source of data, a similar study design would have to be implemented in other regions of the province in order to make proper inference on the concussion rates in these regions. We hypothesize that concussion rates differ between regions meaning data must be collected from regions other than the PCAHA. The current study design would prove too costly to expand to other regions. The use of a single 'gold standard' would be desirable so that data could be collected from volunteers all over the province and all regional estimates could be compared with the one standard. Even more desirable would be to collect the data for the 'gold standard' from within the PCAHA to limit travel costs.

If we assume that all volunteers in the province under report, on average, the same proportion of concussions that actually occurred in their presence, comparison of rates between regions can be performed on volunteer data alone using a simple log-linear model without longitudinal extensions. However, to get proper estimates, an expansion factor could be estimated from data collected within the PCAHA to determine the proportion of concussions being reported by volunteers. This estimate can be used to adjust volunteer reports from other regions. In doing so, we can accumulate data from all over the province and adjust the estimates accordingly providing a more accurate value.

5.4 Computing the Multiplier Adjustment

Let their be two, independent samples of team-games in the PCAHA each with concussion rate λ. One set is sampled by the new 'gold standard' treatment and the other set is sampled by parent volunteers. We will assume that the 'gold standard' will report every concussion that occurs in their observed team-games and the volunteers will only report a portion of theirs. Let X, the number of concussions in a team-game reported by the new 'gold standard', be distributed Poisson(λ). Then let Y, the number of concussions in a team-game reported by the volunteers, be distributed Poisson($p\lambda$) where p is the proportion of concussions reported by the volunteers. Both X and Y are independent. Therefore, $\bar{X}$ and $\bar{Y}$ are the estimated concussion rates for the gold standard and volunteers respectively. The multiplier will be estimated as $\frac{\bar{X}}{\bar{Y}}$.

In order for the multiplier to be an effective tool, its estimate must be precise. Being a ratio of two random variables, the variance of the multiplier can be approximated using Taylor Series expansion of its expected value around its estimate to the first order terms and then taking its expectation (see Mood et al. (1974) pg.181). The variance of the multiplier is approximately

$$\begin{aligned} Var\left(\frac{\bar{X}}{\bar{Y}}\right) &\approx Var(\bar{X})\left(\frac{1}{p\lambda}\right)^2 + Var(\bar{Y})\left(\frac{-\lambda}{(p\lambda)^2}\right)^2 \\ &= \frac{\lambda}{n_{gs}}\left(\frac{1}{p\lambda}\right)^2 + \frac{\lambda}{n_{vol}}\left(\frac{-\lambda}{(p\lambda)^2}\right)^2 \end{aligned}$$

Where n_{gs} and n_{vol} are the number of team-games sampled by the 'gold standard' and the volunteers respectively. The precision will depend on the number of team-games observed by both the 'gold standard' and the parent volunteers, the true concussion rate and the proportion of concussions reported by volunteers (Figure 5.1). The relative errors graphed in Figure 5.1 are based on the minimum and maximum rates estimated from the current data, as well as a much higher rate as suggested by the head of the Motor Behaviour Laboratory (Dr. David Goodman). The value of p was set at 1/3 as it is anticipated that volunteers will not report a proportion lower than this.

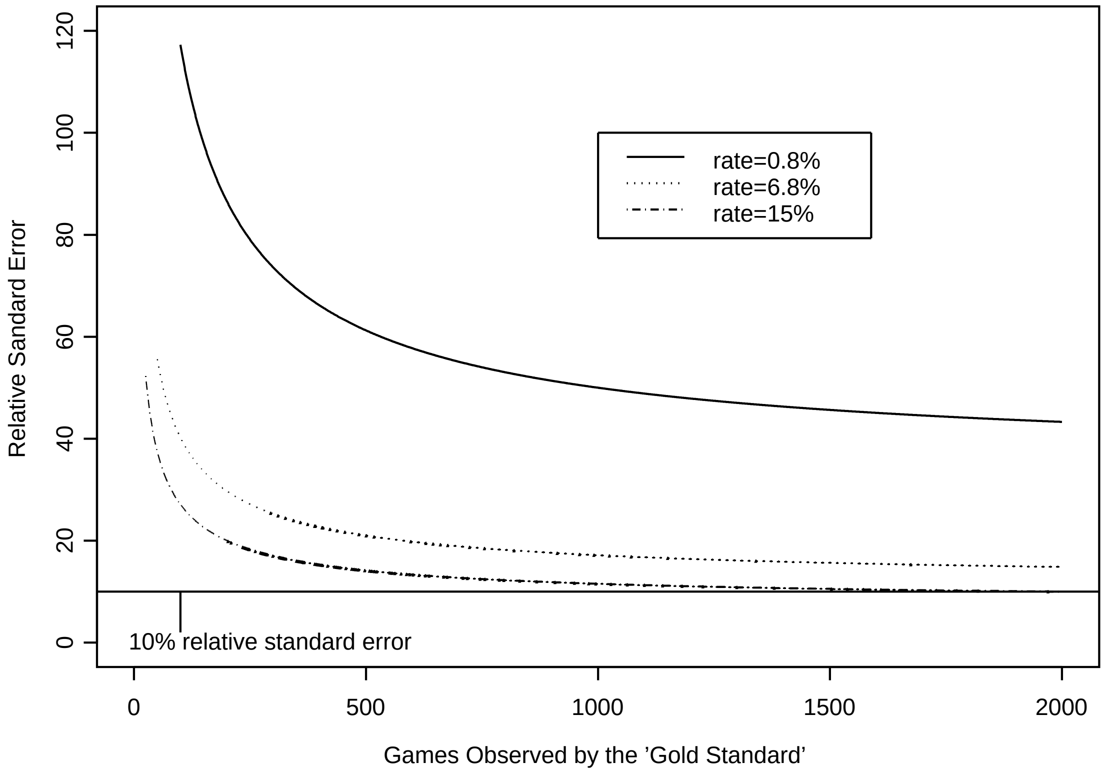


Figure 5.1: Multiplier precision as a function of true concussion rate and sample sizewhile maintaining the number of volunteer sampled team-games at 3000

Even at a very low concussion rate, the slopes in Figure 5.1 begin to level off at around 600 team-games. Also, one must consider that these errors are based on volunteers reporting 1/3 of the actual concussions. If their reporting rate is lower, the curves in Figure 5.1 will be higher. Therefore, it is recommended that 600-800 team-games be observed by the 'gold standard' to account for a wide range of possible concussion rates while remaining conservative with respect to volunteer reporting.

New regions that are added to the study (i.e. Okanagan Mainline) will have data that will be solely collected by volunteers. Factors will remain the same as in the pilot study and will account for skill levels, age groups and association size. Data will still be count data but will focus primarily on concussions. The estimates of the concussion rate obtained from these regions will be multiplied by the expansion factor as $Z\hat{\mu}$ where $Z = \frac{\bar{X}}{\bar{Y}}$, is the expansion factor and $\hat{\mu}$ is the volunteer estimate. The variance of the adjusted estimate is approximated using Taylor Series expansion. The variance of $Z\hat{\mu}$ is

$$Var(Z\hat{\mu}) \approx Var(Z)\hat{\mu}^2 + Var(\hat{\mu})Z^2$$

Using this adjustment, one can still compare factors of interest and examine trends in the data.

One must proceed with caution when using the multiplier from year to year. If volunteers participating in the study become quite good at reporting concussions within a season or two, then applying the expansion factor will overestimate the concussion rate. Some of the participants from the current study will be involved in the study again and their data may indicate that volunteers report the majority of actual concussions. Applying this expansion factor to a new region that is composed of all new volunteers who are not yet experienced at data collection may overestimate the concussion rate as well.

Chapter 6

Conclusion

In an attempt to quantify the incident of concern rate in minor hockey in British Columbia, this study recruited parent volunteers to collect data. To assess the ability of these volunteers, a 'gold standard' was used to collect data on a sub-sample of team-games observed by the volunteers. The rates reported by the volunteers were then compare to the rates provided by the 'gold standard'.

Due to multiple observations on team-games, the analysis used a log-linear model, formed using GEE's. The joint distribution is difficult to formulate making maximum likelihood methods difficult to implement. However, since incidents of concern are counts, the marginal distribution can be assumed to be Poisson-like. It is easier then, to make the assumption that the variance of the joint distribution is a function of its mean and take advantage of the quasi-likelihood methods utilized by GEE's.

The original 'gold standard' was concluded to be insufficient in measuring incidents of concern due to the possibility that each observer (trained or volunteer) was reporting something such as only concussions or all incidents of concern. The term incidents of concern is too subjective and allows for different views of the game to report different events. Therefore, it was not possible to determine whether or not parent volunteers could provide us with reliable data. This led to the need for another study to be conducted that would provide more definitive results.

I feel that the proposed improvements for a future study will result in a clearer determination as to whether parent volunteers are an adequate source of reporting.

The estimation of an expansion factor will allow the adequacy of volunteers to be determined while keeping a lower cost when expanding to other regions of British Columbia. There is a possibility that if the expansion factor is close to one, and volunteers are deemed reliable in terms of reporting, then we can proceed in future seasons to use only the volunteers and will not need to update the expansion factor (the most costly part of the study).

Problems that still remain are under-reporting of concussions and lack of player follow-up. Since players are not assessed after each game, possible concussions may not be reported due to mild or unrecognizable symptoms. Full player assessment would provide the most reliable data. This could only be accomplished by assigning a qualified trainer or physician to each participating team. This would be an ideal, yet extremely costly method of determining concussion rates.

It is my aim that this project contributes to the facilitation of future steps to reduce the rate of concussions throughout minor hockey in not only British Columbia, but across Canada. I feel that in order for players, parents and coaches to be aware of concussions, they must be presented with evidence that the concussion rate may be higher than they are willing to accept as an inherewnt risk of playing hockey. The methods presented in this project will be used to collect data in a cost effective manner. Eventually, using the proposed methods, trends in the concussion rate over time can be examined in hopes that these trends will be declining.

Bibliography

[1] Agresti, A. (2002). "Categorical Data Analysis" 2^{nd} ed. John Wiley and Sons, USA.

[2] Burnham, K. P. and Anderson, D. R. (2002). "Model Selection and Multimodel Inference A Practical Information-Theoretic Approach" 2^{nd} ed. Springer-Verlag, USA.

[3] Dobson, J. A. (1990). "An Introduction to Generalized Linear Models" Chapman and Hall/CRC, USA.

[4] Goodman, D., Gaetz, M. and Meichenbaum, D. (2001). "Concussions in Hockey: There is Cause for Concern" *Medicine and Science in Sports and Excercise*, **33**, 2004-2009.

[5] Hardin, J. W. and Hilbe, J. M. (2003). "Generalized Estimating Equations" Chapman and Hall/CRC, USA.

[6] Liang, K.-Y. and Zeger, S. L. (1986). "Longitudinal Data Analysis using Generalized Linear Models", *Biometrika*, **73**, 13-22.

[7] Lohr, S. L. (1999). "Sampling: Design and Analysis" Duxbury Press, USA.

[8] Mood, A. M. Graybill, F. A. and Boes, D. C. (1974). "Introduction to the Theory of Statistics" 3^{rd} ed. McGraw-Hill, USA.

[9] McCullagh, P. (1983). "Quasi-likelihood Functions", *Annals of Statistics*, **11**, 59-67.

[10] Press Release (2003). "Hockey Season is Here! Hockey Canada Week to Celebrate Canada's Game from Nov 18-15,2003", retrieved from www.hockeycanada.ca/e/news/2003/nr086.html on 16/12/2003.

[11] SAS Institute Inc. (1999). "SAS/STAT® User's Guide, Version 8", Cary, NC:SAS Institute Inc.

[12] Wedderburn, R. W. M. (1974). "Quasi-Likelihood Functions, Generalized Linear Models and the Gauss-Newton Method", *Biometrika*, **61**, 439-447.

[13] Wennberg, R. A. and Tator, C. H. (2003). "National Hockey League Reported Concussions, 1986-1987", *Canadian Journal of Neurological Sciences*, **30**, 206-209.