A GUIDE

TO THE NCAA LACROSSE TOURNAMENT SELECTION PROCESS

1. INTRODUCTION

The culmination of college lacrosse season is the NCAA Championship Tournament. Just as it is for other NCAA varsity sports, there is a selection process where only a limited number of teams are invited to play in the tournament. This guide provides an explanation of how the selection process works for all three men and women's divisions. It also correlates the past ten year's results with selection criteria to produce simple algorithms that predict a team's probability of receiving an invitation to this championship tournament.

1.a. **Committee and Guidelines**

Each division appoints a committee of individuals who select which teams receive an invitation. There are a fixed number of teams invited. Teams are selected either as 'at-large' or for some divisions, 'automatic qualifiers.' The automatic qualifiers or 'AQs' are designated for teams that win certain conferences (usually winners of a playoff tournament) and the 'at-large' selections are made from a pool of all remaining eligible teams. A team typically must have a 0.500 or better record to be considered in that pool. The NCAA provides each committee, a set of guidelines for selecting the at-large teams and these guidelines take the form of primary and secondary criteria. Each division has its own set of criteria as well as number of AQ's and at-large teams. The committees base their selections on the primary criteria and use secondary criteria to separate teams that are initially evaluated to be too close, especially for the final invites. There may be additional discretion by each committee's members in selecting teams. Besides selecting the teams invited, the committee may seed some of the teams and then assign the remaining invitees to first round games against these seeds.

1.b. **Primary and Secondary Criteria**

Primary and Secondary criteria for Men and Women's Division I is based on something called the Rating percentage Index or RPI. A component of the RPI is known as the Strength of Schedule or SOS. In addition, for men's Division I, wins/losses against teams ranked by RPI is known as Quality Wins. RPI, SOS and Quality wins is explored below. For men and women's division II and III, W-L records and not RPI are the primary criteria. However, SOS which is derived when computing RPI is a factor here. Strength of Schedule is considered against all teams, regional teams and divisional teams. The win-loss record against all teams, regional teams, out-of-region teams and divisional teams are employed. For women's division I and II,

performance in the last six games and wins against teams with a record greater than 0.75 serve as additional criteria.

1.c. Ratings Percentage Index (RPI)

RPI (ratings percentage index) is a rating that applies to a team's record and schedule such that the stronger the opponents, the higher the RPI rating. Conversely, the weaker the opponents, the lower the RPI rating. In order to tabulate an RPI rating, each opponent's strength has to be evaluated by examining its W-L record as well as the W-L record of its opponents.

*RPI = (1/4 * team's record) + (1/2 * opponents' record) + (1/4 * opponents' opponents' record)*

Records are in the usual form of a win percentage: wins / (wins + losses). Ties are treated as 1/2 win + 1/2 loss. Here is an example of the computation:

A team has played 8 of its 12 games through the season. The team's record is 7 wins and 1 loss or 0.875.

The records for the opponents are as follows:

```
Team A        5   3   .625
Team B        4   3   .571
Team C        2   5   .286
Team D        3   4   .428
Team E        7   1   .875
Team F        2   5   .286
Team G        2   5   .286
Team H        9   0  1.000
             --  --
             34  26   .545
```

Note that the .545 is not based on 34/60 (.567) but on the average of each opponent's percentage in the right-hand column. Games against the team under consideration is not counted.

The records for each opponent's opponents are

```
Team A       42  18   .750  - based on average of all of a team's opponents
Team B       22  24   .502    and not based on total wins / total games
Team C       18  18   .483
Team D       32  22   .600
Team E       30  31   .530
Team F       29  15   .624
Team G       37  16   .645
Team H       39  27   .615
            ---  --
            249 171   .594
```

Note that the .594 opponent-opponent figure above is not based on 249/(249+171) but rather is the average of the percentages in the last column. Games against the opponent under

consideration is not counted.

The RPI for our example team is then computed as follows:

Eqn (1) RPI = (0.25 * 0.875) + (0.50 * 0.545) + (0.25 * 0.594) = 0.600

The RPI is a staple for selections in other varsity sports, most notably basketball and hockey. The method has also come under great scrutiny and criticism because of its limitations. For example, only wins and losses count and goal or point margins are not considered. That means a game where a team that wins by 10 or more goals counts no differently than a game won by a team by one goal in overtime. Second, the site of the game is ignored meaning a road victory and a home victory (and loss) are counted equally despite a home-field advantage. Next as in all methods, the sample size (number of games) may be too small leading to erroneous conclusions. As in all algorithms, the method breaks down if there is not enough cross-over play between leagues or conferences. For example, if all teams in Alaska played only teams in Alaska and all teams in Hawaii played only teams in Hawaii and the RPI was used for combining both states, the teams with the best record in each state would have the highest RPI rankings. However, if one state was much better than another state, the data would not reflect this difference. On the other hand if teams from Hawaii played teams from Alaska, then the results probably would reflect that teams from one state are stronger than teams from another state because the win-loss records would change due to the out-of-state games. This would occur only if there were sufficient games played between these states. There are modifications to the RPI to take into account some of these issues so that hockey's RPI method differs from the lacrosse RPI method.

1.d. Strength of Schedule (SOS)

Strength of Schedule (SOS) is a rating which applies to a team's schedule such that the stronger the opponents, the higher the SOS rating, or conversely, the weaker the opponents, the lower the SOS rating. In order to tabulate an SOS rating, the opponent's strength has to be evaluated.

The RPI method, discussed above, is based on a team's record and its opponents' records and their opponents' records. The 2nd and 3rd components, when added, yield the strength of schedule. This method is based on wins and losses and not goal or point margins or other factors .

Consider this example of Team A's record and its opponents' and opponents' opponents records:

Team Rec (1/4) W/L PCT.	Opp Rec (1/2) W/L PCT.	Opp-Opp Rec (1/4) W/L PCT.
0.75	0.60	0.40

Using the second and third weights and values shown above, Team A's SOS would be ((0.50 * 0.60) + (0.25 * 0.40)) = 0.40.

The SOS comes under the same criticism as the RPI since it is derived from the RPI. Another issue with the SOS is that there is no mathematical evidence that it truly represents strength of schedule because of the issues mentioned above.

1.e. Quality Wins

Quality Wins is a measure of a team's performance weighted by the strength of an opponent. Sometimes they are defined as 'significant wins.' For men's division I they include not only weighted wins but also weighted losses. As an example, using the rpi rankings as a measure of a teams strength, a team will earn more points defeating a higher ranked team then a lesser ranked teams. The same hold trues for losing, where a loss to a higher ranked team will subtract less points than losing to a lower ranked team.

```
Defeating an RPI ranked team:
          1 -  5 = 25 points added
          6 - 10 = 20 points added
         11 - 20 = 15 points added
            > 20 =  5 points added

Losing to an RPI ranked team:
          1 -  5 =  5 points subtracted
          6 - 10 = 15 points subtracted
         11 - 20 = 20 points subtracted
            > 20 = 25 points subtracted
```

1.f. Win-Loss Percentage

Win Loss percentage is calculated by taking the number of wins and dividing it by the total number games played. A tie is treated as a half win. The games that count may be all games, only those games played in a region, only those games played outside a region and only those games played within a division.

2. What about Polls and Computer Ratings

No where in all of the NCAA criteria for all divisions is there any mention of polls. No outside opinions are considered in the selection process. One of the reasons that there is always 'shock' and disappointment when the selections are announced is that the guidelines for selection can produce results which differ significantly from commonly accepted views, i.e., the polls. There have been occasions where undefeated teams and teams ranked as high as fifth in the polls did not receive an invitation. Outside computer rankings are not employed in the selection process but some committees are provided internal computer results based on internal NCAA algorithms.

3. Developing an Algorithm to Predict Team Selections

3.a. How Was The Algorithms Developed For Each Division?

The algorithm or formula for each division has two objectives: (1) to quantify the primary and secondary criteria into a numerical expression that (2) accurately predicts those teams that will be selected and those teams that will not be selected. For some of the components such as RPI, SOS and win-loss records, this was easy. For other criteria such as head-to-head performance or comparison between common opponents, this was difficult. In order to get around this, an assumption was made that head-to-head and common opponents only play a role when two teams are close and since the goal is to predict probabilities, applying these tie-breaking criteria was not necessary. The next issue was assigning a weight factor to these components. The algorithm simply weighted secondary criteria as ½ of primary criteria. Most components of the formula were based on rankings as opposed to actual rating since it was assumed that might more closely simulate the selection process. In the case of the two women's division two special components, absolute values instead of rankings were used. There are a multitude of formulas that could be derived, the ones described below were selected because they results in clear regions of when a team was a "lock", "in", "bubble in and out" and "out".

3.b. Evaluating the six divisions

The remainder of this guide describes each division in some detail. Initially, the total number of teams selected, total number of automatic qualifiers and total number of at-large teams is listed. This is followed by a list of the primary and secondary criteria. The algorithm or formula used to predict the probabilities is then listed followed by a discussion of a histogram representing over the past ten years, how often a team with a certain sum of the formula either was invited to the tournament or rejected by the committee.

3.c. Determining Automatic Qualifier Probabilities

The probability of a team receiving an invitation is based on the probability of each team in a designated AQ conference of (1) being selected to the league's 4-team post season tournament and (2) winning that tournament. The numerical method takes the current league win-loss record and performs a 100,000 simulations where the remaining unplayed games are assigned values based on the power ratings of the opposing teams and a random number generator. The best four teams than play in a simulated post season tournament and the winner is computed. Based on these 100,000 simulations, the number of times a team wins out of 100,000 represents the probability (*100) of getting an AQ invitation.

4.a. MEN'S DIVISION I

- Total Number of Teams Selected: 16
- Total Number of Automatic Qualifier Teams: 7

  American East Conference
  Big East Conference
  Colonial Athletic Conference
  Eastern College Athletic Conference
  Ivy Group
  Metro Atlantic Athletic Conference
  Patriot League

- Total Number of At-large Teams: 9

Primary Criteria:

- Strength of schedule Index based on the 10 highest-ranking opponents in RPI. Two games against the same opponent count as two contests.
- Results of the RPI
  - Record against teams ranked 1-5, 6-10, 11-20, and 21+.
  - Average RPI win (average RPI of all wins)
  - Average RPI loss (average RPI of all losses)
- Head-to-Head Competition
  - Results against common opponents
  - Significant wins (against teams ranked higher in the RPI)
  - Significant losses (against teams ranked lower in the RPI)
  - Locations of contest

Secondary Criteria: None

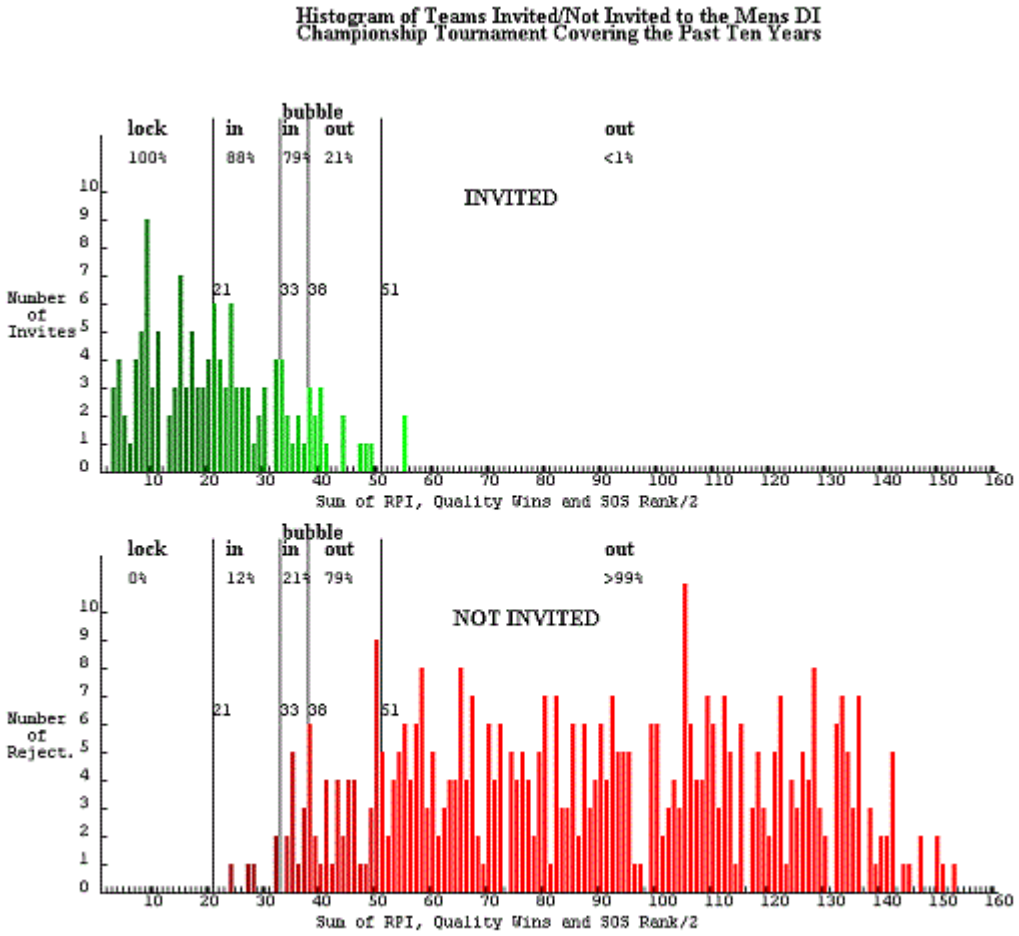
Formula: the sum of

- RPI ranking +
- QW ranking +
- ½ SOS ranking

Examining the histogram below, the independent coordinate is the sum of the rpi ranking, the quality wins ranking and ½ the strength of schedule ranking. The last three automatic qualifiers were excluded because their sums were skewed to the right. Teams selected to the tournament over the last ten years had a sum lying between 1 and 53. All teams with a sum less than 21 received an invitation 100% of the time. Teams with a sum between 22 and 33 received an

invitation 88% of the time. Teams with a sum between 34 and 51 lied in the bubble with 34-38 yielding a 79% invitation and 39-51 yielding a 21% invitation. Teams with a greater than 51 sum received an invitation less than 1% of the time.

Figure 1. Histogram of Results for Men's Division I from 2003 to 2012.



Lock	In	Bubble In	Bubble Out	Out
1-21 (100%)	22-33 (92%)	34-38 (79%)	39-51 (21%)	>51 (<1%)

4.b. MEN'S DIVISION II

- Total Number of Teams Selected: 4
- Total Number of Regional Automatic Qualifier Teams: 3
- Total Number of At-large Teams: 1

Primary Criteria:

- Overall won-lost record
- In-region won-loss record
- Overall strength of schedule
- Head-to-head competition
- Results against common opponents

Secondary Criteria:

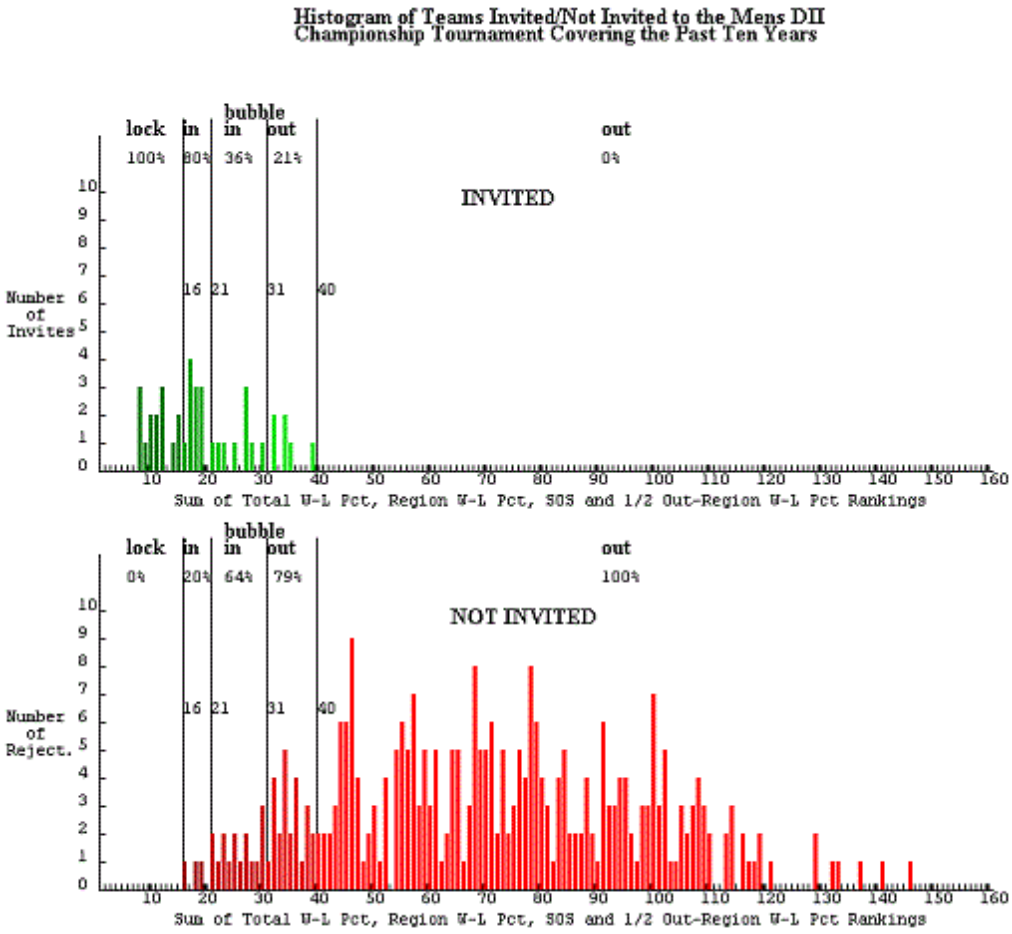
- Results against Division II out-of-region competition
- Results against teams already selected

Formula: the sum of

- Total Win-Loss ranking +
- Regional win-loss ranking +
- SOS ranking +
- Out-of-Region Win-Loss ranking

Examining the histogram below, the independent coordinate is the sum of the total win-loss ranking, the regional win-loss ranking, the sos ranking and ½ the out-of-region win-loss ranking. Teams selected to the tournament over the last ten years had a sum lying between 1 and 40. All teams with a sum less than 16 received an invitation 100% of the time. Teams with a sum between 17 and 21 received an invitation 80% of the time. Teams with a sum between 22 and 40 lied in the bubble with 22-31 yielding a 36% invitation and 32-40 yielding a 21% invitation. Teams with a greater than 40 sum received no invitation.

Figure 2. Histogram of Results for Men's Division II from 2003 to 2012.



Lock	In	Bubble In	Bubble Out	Out
1-16 (100%)	17-21 (80%)	22-31 (36%)	32-40 (21%)	>40 (0%)

4.c. MEN'S DIVISION III

- Total Number of Teams Selected: 28
- Total Number of Regional Automatic Qualifier Teams: 18

Commonwealth Coast
Empire 8
Great Northeast
Liberty
Little East

NESCAC
North Atlantic
Pilgrim
Skyline
SUNYAC

Capital
Centennial
Colonial States
Landmark
Middle Atlantic
Midwest
ODAC
SCAC

- Total Number of At-large Teams: 10

Primary Criteria

- Win-loss percentage against regional opponents
- Strength of schedule vs. regional opponents
- In-region head-to-head competition
- In-region results vs. common regional opponents
- In-region results vs. regionally ranked teams

Secondary Criteria

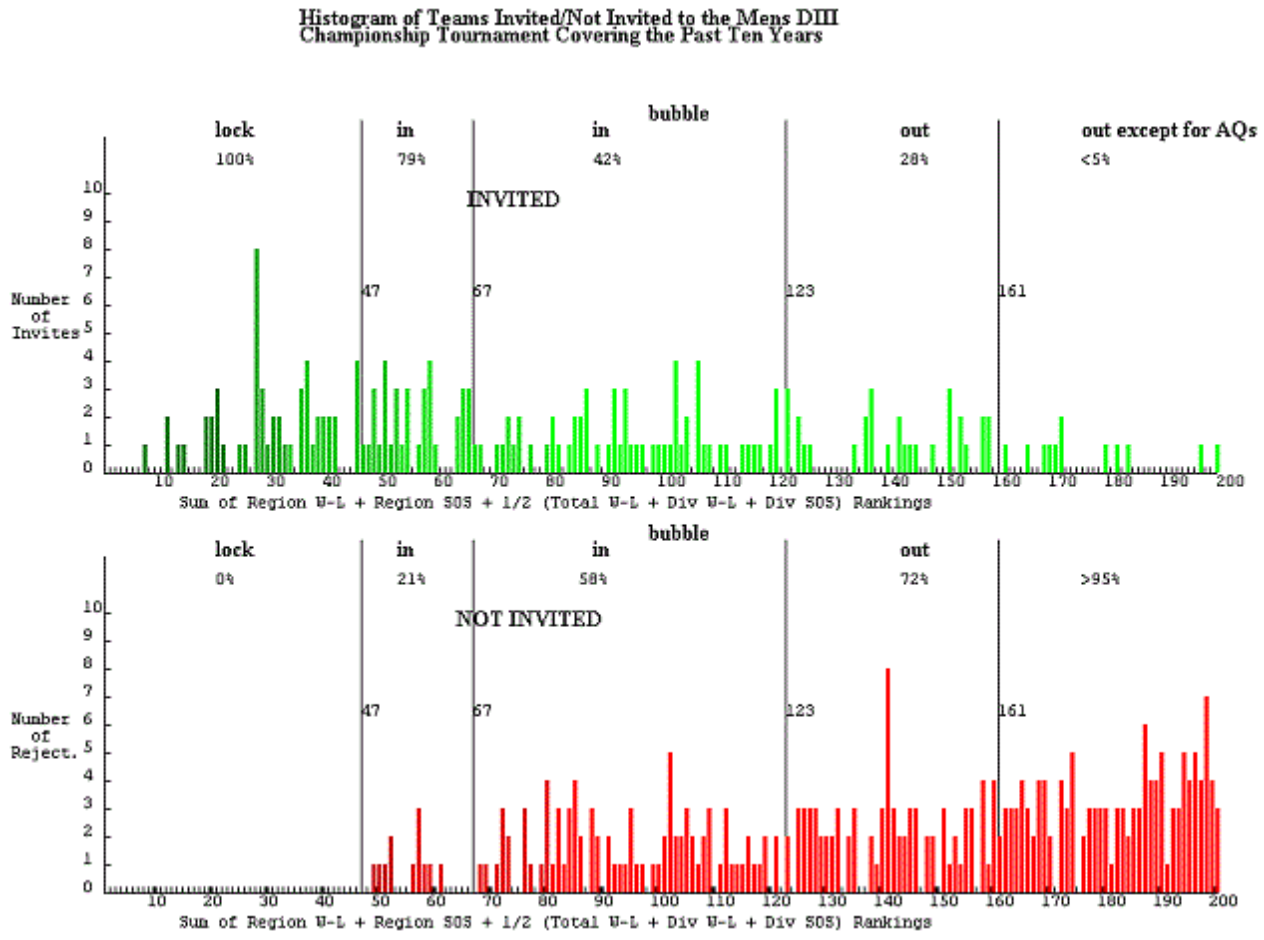
- Out-of-region head-to-head competition
- Overall Division III won-loss percentage
- Results vs. common non-Division III opponents
- Results vs. all Division III ranked teams
- Overall win-loss percentage
- Results vs. all common opponents
- Overall Division III strength of schedule

Formula: the sum of

- Regional Win-Loss ranking +
- regional SOS ranking +
- ½ Divisional Win-Loss ranking +
- ½ Regional SOS ranking +
- ½ Division SOS ranking

Examining the histogram below, the independent coordinate is the sum of the regional win-loss ranking, the regional sos ranking and ½ the total won-loss ranking + ½ the divisional won-loss ranking + ½ the divisional sos ranking . Teams selected to the tournament over the last ten years had a sum lying between 1 and 200. All teams with a sum less than 47 received an invitation 100% of the time. Teams with a sum between 48 and 67 received an invitation 79% of the time. Teams with a sum between 68 and 161 lied in the bubble with 68-123 yielding a 42% invitation and 124-161 yielding a 28% invitation. Teams with a greater than 161 sum received an invitation less than 5% of the time.

Figure 3. Histogram of Results for Men's Division III from 2003 to 2012.



Lock In Bubble In Bubble Out Out

1-47 (100%)	48-67 (79%)	68-123 (42%)	124-161 (28%)	>161 (<5%)

4.d. WOMEN'S DIVISION I

- Total Number of Teams Selected: 26
- Total Number of Regional Automatic Qualifier Teams: 13

  Amer. Lac. Conf.
  America East
  Atlantic 10
  Atlantic Coast
  Atlantic Sun
  Big East
  Big South
  Colonial
  Ivy League
  Metro Atlantic
  MPSF
  Northeast
  Patriot League

- Total Number of At-large Teams: 13

Primary Criteria:

- Rating Percentage Index (RPI)
- Head-to-head competition
- Results vs. common opponents
- Significant wins and losses
- Evaluation of the 10 highest-rated teams on an institutions schedule (defined by SOS and winning percentage)

Secondary Criteria:

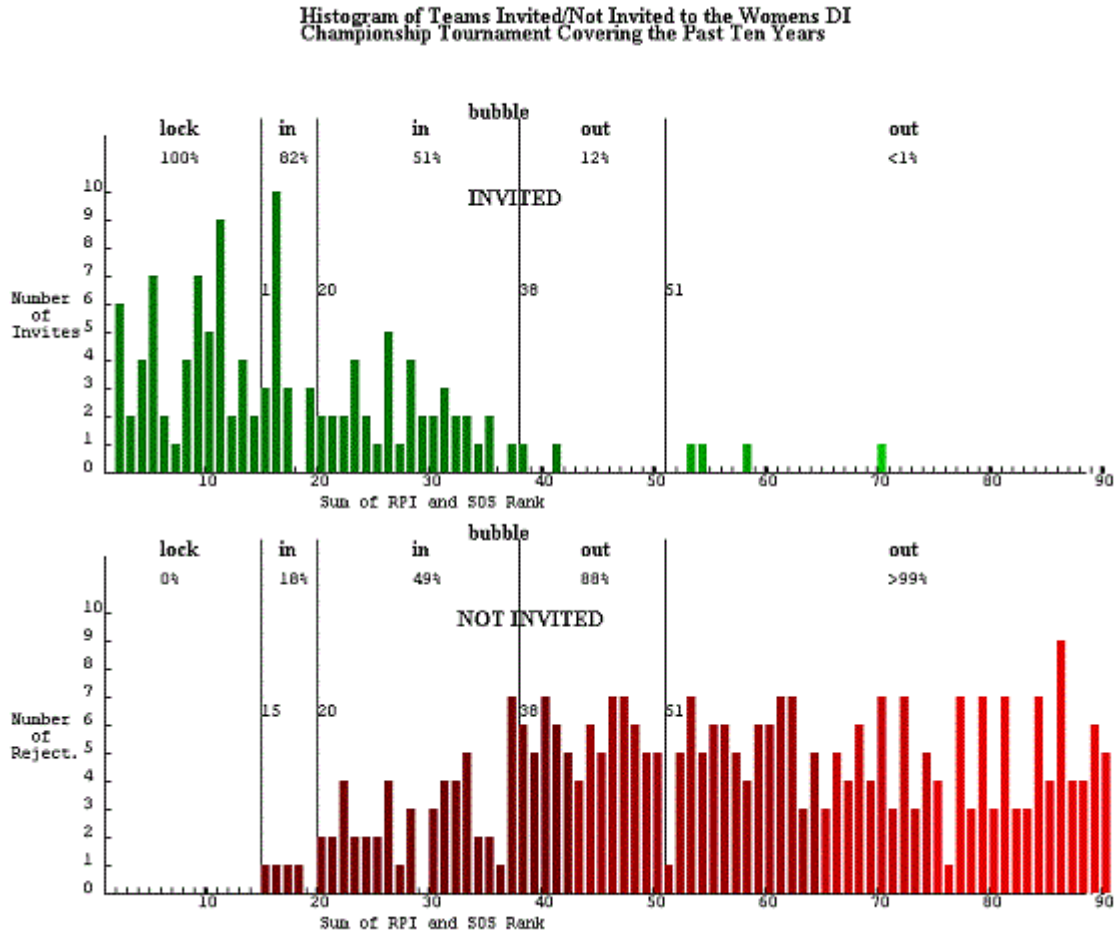
- Late season performance in last six games

Formula: the sum of

- RPI ranking +
- SOS ranking

Examining the histogram below, the independent coordinate is the sum of the rpi ranking and sos ranking. Teams selected to the tournament over the last ten years had a sum lying between 1 and 51. All teams with a sum less than 15 received an invitation 100% of the time. Teams with a sum between 16 and 20 received an invitation 82% of the time. Teams with a sum between 21 and 51 lied in the bubble with 21-38 yielding a 51% invitation and 39-51 yielding a 12% invitation. Teams with a greater than 51 sum received an invitation less than 1% of the time.

Figure 4. Histogram of Results for Women's Division I from 2003 to 2012.



Lock	In	Bubble In	Bubble Out	Out
1-15	16-20	21-38	39-51	>51

(100%)	(82%)	(51%)	(12%)	(<1%)

4.e. WOMEN'S DIVISION II

- Total Number of Teams Selected: 6
- Total Number of Regional Automatic Qualifier Teams: 3 per region
- Total Number of At-large Teams: 0

Primary Criteria

- Division II in-region won-lost record
- Strength of schedule in-region
- Results of head-to-head competition
- Results vs. common opponents
- Wins over teams with records above .750
- Overall Division II record

Secondary Criteria

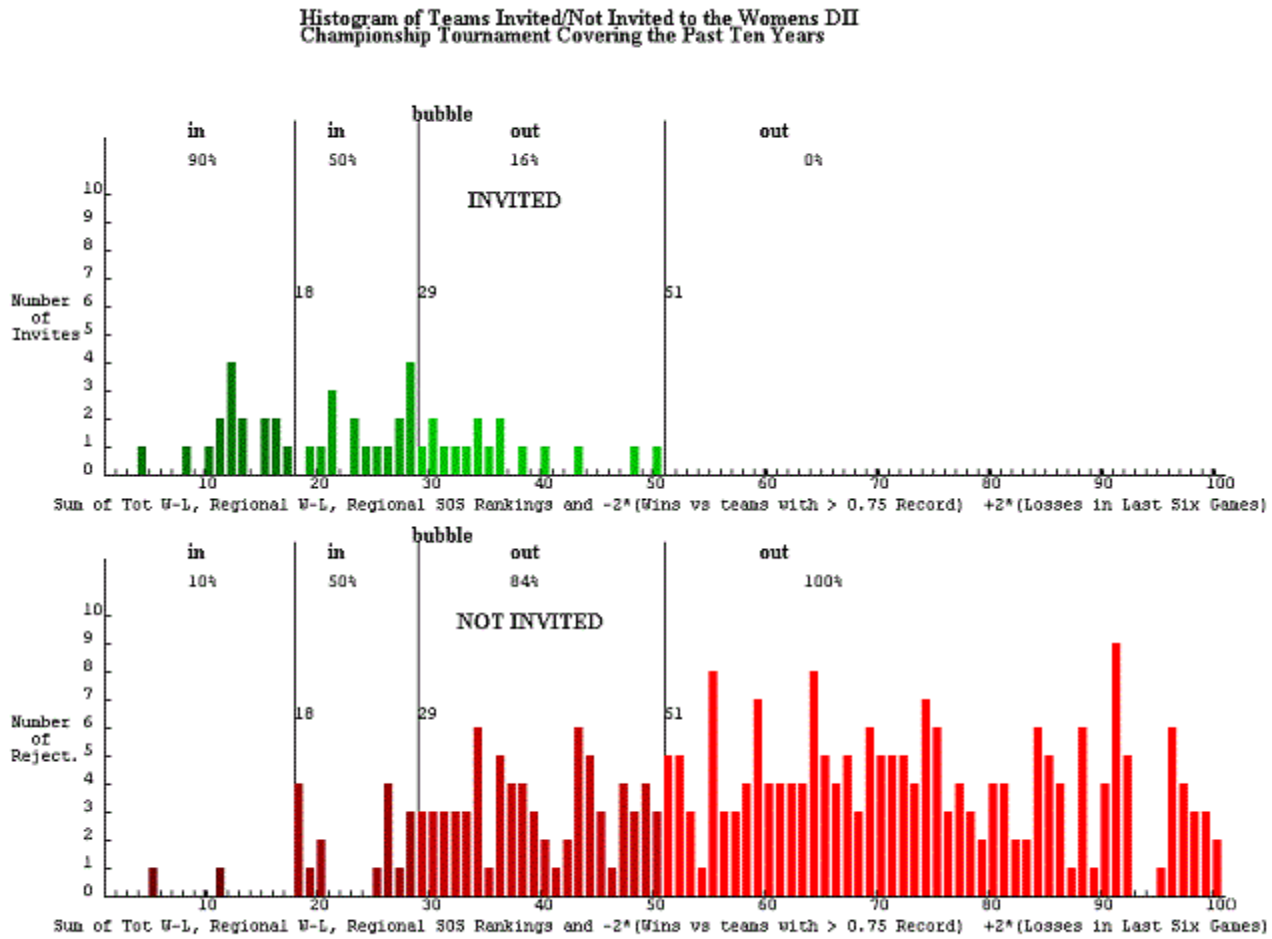
- Late season performance can be employed as a secondary selection criterion.

Formula: the sum of

- Total Win-Loss ranking +
- Regional Win-Loss ranking +
- 2*Losses in Last Six Games +
- Regional SOS ranking –
- 2* Wins versus teams with a record > 0.75

Examining the histogram below, the independent coordinate is the sum of the total win-loss ranking, the regional win-loss ranking, the regional sos ranking minus 2 * the number of wins over teams with a win-loss record greater than 0.75 and – 2 * the number of losses in a team's last six games . Teams selected to the tournament over the last ten years had a sum lying between 1 and 51. All teams with a sum less than 18 received an invitation 90% of the time. Teams with a sum between 19 and 51 lied in the bubble with 19-29 yielding a 50% invitation and 30-51 yielding a 16% invitation. Teams with a greater than 51 sum did not receive an invitation.

Figure 5. Histogram of Results for Women's Division II from 2003 to 2012.



Lock	In	Bubble In	Bubble Out	Out
N/A	1-18 (90%)	19-29 (50%)	30-51 (16%)	>51 (<0%)

4.f. WOMEN'S DIVISION III

- Total Number of Teams Selected: 31
- Total Number of Regional Automatic Qualifier Teams: 18

Capital Athletic Conference
Centennial Conference
Colonial States Athletic Conference
Commonwealth Coast Conference
Commonwealth Conference
Empire 8 Conference

Freedom Conference
Great Northeast Athletic Conference
Liberty League
Little East Conference
New England Small College Athletic Conference
New England Women's and Men's Athletic Conference
New England Women's Alliance
North Coast Athletic Conference
Old Dominion Athletic Conference
Skyline Conference
State University of New York Athletic Conference
USA South Athletic Conference

- Total Number of At-large Teams: 13

Primary Criteria:

- Win-loss percentage against regional opponents
- Strength of schedule
- In-region head-to-head competition
- In-region results vs. common regional opponents

Secondary Criteria:

- Overall Division III strength of schedule
- Division III opponents' average winning percentage
- Division III opponent's opponent's average winning percentage
- Out-of-region head-to-head competition
- Overall Division III win-loss percentage
- Results vs. common non-Division III opponents
- Results vs. Division III teams ranked in other regions
- Overall win-loss percentage
- Results vs. common out-of-region opponents

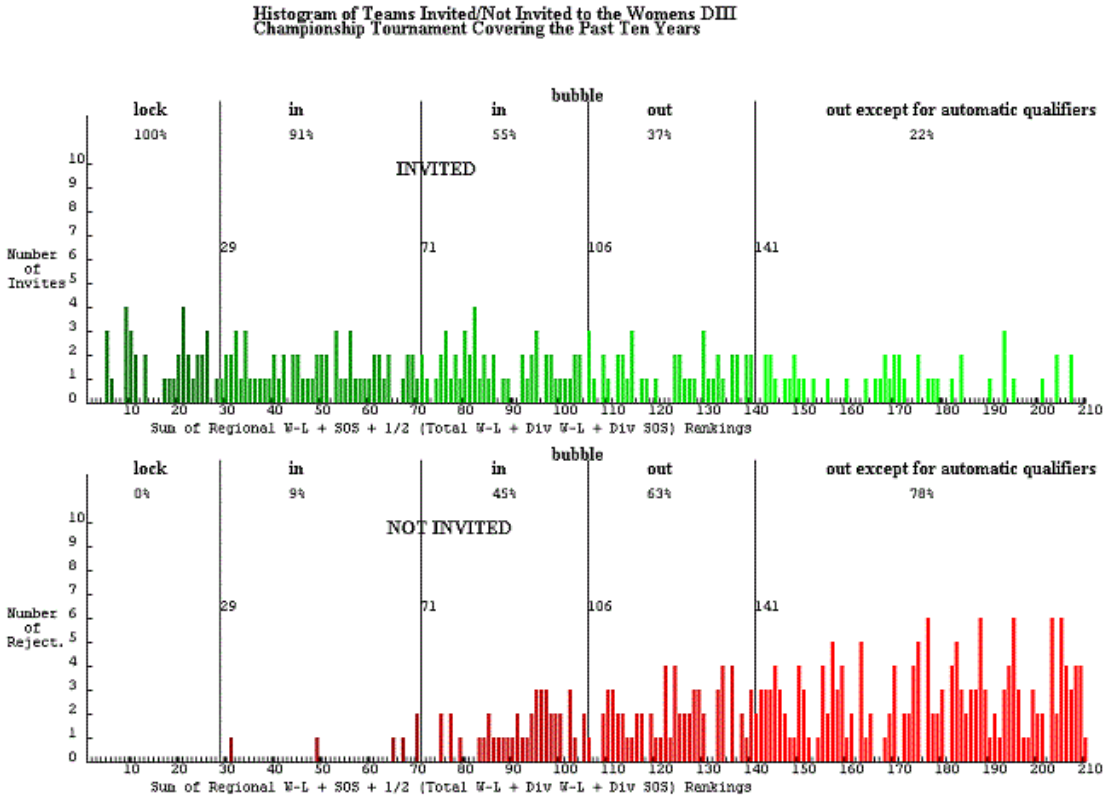
Formula: the sum of

- Total Win-Loss ranking +
- Regional SOS ranking +
- ½ Total Win-Loss ranking +
- ½ Divisional Win-Loss ranking +
- ½ Divisional SOS ranking

Examining the histogram below, the independent coordinate is the sum of the regional win-loss ranking, the sos ranking and ½ the total won-loss ranking + ½ the divisional won-loss ranking +

½ the divisional sos ranking . Teams selected to the tournament over the last ten years had a sum lying between 1 and 210. All teams with a sum less than 29 received an invitation 100% of the time. Teams with a sum between 30 and 71 received an invitation 91% of the time. Teams with a sum between 72 and 141 lied in the bubble with 71-106 yielding a 55% invitation and 107-141 yielding a 37% invitation. Teams with a greater than 141 sum received an invitation less than 22% of the time.

Figure 6. Histogram of Results for Women's Division I from 2003 to 2012.



Results vs common out-of-region opponents

Lock	In	Bubble In	Bubble Out	Out
1-29 (100%)	30-71 (91%)	72-106 (55%)	107-141 (37%)	>141 (22%)

5. CONCLUSIONS

In conclusion, based on the discussions above, a summation of all the algorithms for each division and gender is shown below. The actual computations and probabilities appear throughout the season at laxpower.com. The following table quantifies the primary and secondary factors of the NCAA Selection Criteria for a team receiving an invitation to the NCAA Championship Tournament. The formulas and probabilities are for all three divisions of men and women's varsity lacrosse. The components for each division's algorithm is typically based on primary and secondary criteria for that division.

A 'lock' means a team is virtually guaranteed an invitation. 'In' means a team will, in all probability be invited. 'Bubble In' means a team is leaning towards an invitation and 'bubble out' means a team has a greater probability of not getting an invite. 'Out' means a team will in all probability not receive an invitation.

The first set a numbers, eg., "1-21" specifies the range of the sum of the rankings according to the formula and the number below it represents the probability in percent of that team being selected to the tournament, e.g., "100%." As another example, for a team in Men's Division III, if the sum of the rankings for RR + SOSR + ½(RD+SOSR+SOSD) lies between 68 and 123, than, that team is on the 'inside' of the bubble and has a 42% chance of being selected for the tournament based on the last 10 years of data. Another example is if a Women's Division I team has an RPI and SOS total which lies between 15 and 20, then that team has an 82% likelihood of making the tournament based on the fact that over the last ten years, 82% of the teams in that division made the tournament when their sum lies in the 15 to 20 range.

Table 1. Formulas and Predictions For Each Division

Div.	*Formula* (see symbol explanation below).	*Lock*	*In*	*Bubble In*	*Bubble Out*	*Out*
Men's Div I	RPI + QW + ½ SOS	1-21 (100%)	22-33 (92%)	34-38 (79%)	39-51 (21%)	>51 (<1%)
Men's Div II	RT + RR + SOS+RO	1-16 (100%)	16-21 (80%)	21-31 (36%)	31-40 (21%)	>40 (0%)
Men's Div III	RR + SOSR + ½ {RD + SOSR + SOSD}	1-47 (100%)	48-67 (79%)	68-123 (42%)	124-161 (28%)	>161 (<5%)
Women's Div I	RPI + SOS	1-15 (100%)	15-20 (82%)	20-38 (51%)	38-51 (12%)	>51 (<1%)
Women's Div II	RT + RR + 2*RL + SOSR - 2*W75	N/A	1-18 (90%)	19-29 (50%)	30-51 (16%)	>51 (<0%)
Women's Div III	RT +SOSR + ½ {RT + RD + SOSD}	1-29 (100%)	29-70 (91%)	72-106 (55%)	107-141 (37%)	>141 (22%)

(x-y) = range of sum; z% = per cent of teams within that range that made the tournament

Table 2. Symbols for Table I.

RPI	Ranking of Rating Percentage Index
QW	Ranking of Quality Wins
SOS	Ranking of Strength of Schedule
RT	Ranking of (W-L) for all games
RD	Ranking of (W-L) for all division games
RR	Ranking of (W-L) for all regional games
RO	Ranking of (W-L) for all out-of-regional games
RL	Number of Losses in last six games
SOSD	Ranking of Strength of Schedule for divisional games
SOSR	Ranking of Strength of Schedule for regional games
W75	Number of Wins against teams with Record > 0.75