

National Conference of State Legislatures
The Future of Elections
Williamsburg, VA – June 15, 2015

Risk-limiting Audits in Colorado

Dwight Shellman
County Support Manager
Colorado Department of State, Elections Division

Context – Colorado's Election Model

- **Mail ballots** – every active elector automatically receives a mail ballot
- **Same-day registration** – every eligible individual can register to vote until 7 pm on Election Day
- **Vote centers** – every voter can go to any vote center in his or her county and obtain or cast an official (rather than provisional) mail or in-person ballot

Context – Voting Patterns & Processes

- Since Colorado's current election model was adopted in 2013, and depending on the election, **92-97% of total ballots cast are mail ballots**
- For variety of reasons, all 64 counties now **centrally tabulate paper ballots**, including those cast in-person at vote centers
- 54 of 64 counties will transition to new voting system this year, representing 91% of active electorate. New system is capable of exporting ballot-level CVRs (explained below)

Post-election Audits in Colorado

- Before 2017, Colorado law required county clerks to conduct **random post-election audits** after Election Day and before certifying official results
- Starting in 2017, Colorado law requires counties to conduct **risk-limiting audits (RLAs)**
- Secretary of State recently published proposed Election Rules specifying manner in which RLAs must be conducted - available at sos.state.co.us
- Anyone can submit written comments

Random Audit Methodology

- Counties submit to SOS inventory of the voting devices they will use before each election
- For each county, SOS randomly selects at least 1 DRE and 1 ballot scanner that county must audit
- County is required to audit 20% of (but not more than 500) ballots tabulated on randomly selected devices
- For DREs, counties manually tabulate votes on VVPAT and verify hand tally against DRE's tabulation tape

Random Audits (continued)

- For paper ballots, counties must retrieve ballots tabulated on each separate ballot scanner and
 - Reset randomly selected scanner to zero
 - Randomly select minimum number of ballots to audit
 - Re-scan ballots to be audited and print tabulation tape
 - Manually tabulate audited ballots
 - Verify that hand tally of audited ballots matches tabulation tape for those ballots
- Counties file reports of audits with SOS

Random Audits (continued)

- Provides some evidence that results are accurate:
 - Devices accurately tabulated ballots during pre-election logic and accuracy test (LAT)
 - Devices were sealed and chain-of-custody maintained immediately following LAT
 - Humans verify that randomly selected devices accurately tabulated a randomly selected subset of all ballots cast after the election
 - No reason to believe unaudited devices performed any differently on or before Election Day
 - Therefore, we have confidence in the results

Weaknesses of random PEAs

- Audits specific devices, not the election
- Audit provides some evidence but does not itself prove that outcome of election is accurate
- Audit verifies an artificially created subset of results, not the actual results
- Level of confidence yielded by random audit is not statistically meaningful or significant
- But it was the best we could do with the available technology

Risk-limiting audits (RLAs)

- RLAs eliminate the weaknesses of random post-election audits
- RLA provides strong statistical evidence that the election tabulation outcome is right, and has a high probability of discovering and correcting a wrong outcome.
- In other words, RLAs limit the risk that an incorrect election tabulation outcome will escape discovery and correction during the audit.

Notable RLA advocates and literature

- Dr. Philip Stark (University of California at Berkeley) introduced the concept of post-election RLAs in 2008
- Dr. Stark and Dr. Mark Lindeman (Columbia University) published an article (tragically and deceptively) entitled “*A Gentle Introduction to Risk-Limiting Audits*” in 2012
- Ron Rivest (Massachusetts Institute of Technology) developed a pseudo-random number generator ensuring that the audited ballots truly represent a random sampling of all ballots cast in the election
- In Colorado, Harvie Branscomb and Neal McBurnett

RLA Concepts and Terms

- RLA continues until the risk limit is satisfied or a full hand count results
- Risk limit: The largest chance that a wrong outcome will not be discovered and corrected in the audit
- Outcome: The winner and losers, not the exact votes they received
 - Right (or correct) outcome: The reported winning candidate or choice matches the actual winning candidate or voting choice
 - Wrong (or incorrect) outcome: The reported winning candidate or voting choice does not match the actual winning candidate or voting choice

RLA Concepts and Terms (continued)

- If the risk limit is 5% and the outcome is wrong, there is at most a 5% chance that the audit will not discover the error and correct the outcome, and at least a 95% chance that the audit will correct the outcome.
- The number of ballots that must be audited to satisfy the risk limit is based on the smallest margin of the contest selected for audit and the risk limit.
- The smaller the margin, the more ballots to audit. The smaller the risk limit, the more ballots to audit.

RLA Concepts and Terms (continued)

- The initial sample size in a comparison audit RLA is determined by the following algorithm:

$$-2g \log(a) / ((m + 2g(r_1 \log(1 - 1/(2g)) + r_2 \log(1 - 1/g) + s_1 \log(1 + 1/(2g)) + s_2 \log(1 + 1/g))))$$

- This formula tells us when the comparison audit can stop:

$$-2g(\log(a) + o_1 \log(1 - 1/(2g)) + o_2 \log(1 - 1/g) + u_1 \log(1 + 1/(2g)) + u_2 \log(1 + 1/g)) / m)$$

RLA Concepts and Terms (continued)

- Ballot-level cast vote record (CVRs): Data file that shows how the voting system interpreted voter markings on each individual ballot
- Each CVR in the file has a unique identifier
- In Colorado, the unique identifier is a concatenation of:
[Scanner ID #] + [Batch #] + [Ballot's position within batch]
- To associate a tabulated paper ballot with corresponding CVR, counties must either imprint CVR number on ballot during tabulation, or segregate tabulated ballots by device on which they are scanned, and maintain ballots within batches in the same order they are scanned

RLA Concepts and Terms (continued)

- Sample CVR

[illegible]

RLA Concepts and Terms (continued)

- Ballot manifest: A record maintained by county that functions as its road map during the RLA
- Sample ballot manifest:

COUNTY	SCANNER	BATCH	BALLOTS	BIN
Fremont	03	01	49	17
Fremont	03	02	50	17
Fremont	03	03	52	18
Fremont	03	04	48	18

RLA Workload Examples

2016 Presidential Contest	
Total Ballots Cast = 2,859,216; Risk Limit = 5%	
Clinton	1,338,870
Trump	1,202,484
Smallest Margin = 136,386 Diluted Margin = smallest margin/total ballots cast = 4.77%	
Using Dr. Stark's comparison RLA algorithm, the number of ballots to audit is 142 for the whole state. (In our current audit, all counties are probably required to audit at least 32,000 ballots)	

RLA Workload Examples (continued)

City and County of Denver Referred Question 2A	
Total Ballots Cast = 341,987; Risk Limit = 5%	
Yes/For	235,595
No/Against	75,598
Smallest Margin = 159,997 Diluted Margin = smallest margin/total ballots cast = 46.8%	
Using Dr. Stark's comparison RLA algorithm, the number of ballots to audit is 15 for the county. (In our current audit, Denver is probably required to audit at least 500 ballots)	

Implementing RLAs

- We are developing a RLA software tool for counties and state to use to conduct RLAs
- RLA Tool will consist of several features
 - Philip Stark's algorithms to calculate the number of ballots based on the risk limit established by Secretary Williams and the margin of contests
 - Ron Rivest's pseudo-random number generator, to randomly select ballots to audit
 - County UI to upload ballot manifests, CVRs, and summary results
- Once developed, RLA Tool will be open source and available to anyone who wants it without charge

Implementing RLAs (continued)

- Secretary of State will establish the risk limit and select at least one statewide contest and one countywide contest in each county for audit based on closeness of margins, number of eligible electors, workload on counties, other factors
- On 9th day after election, counties will use RLA Tool to upload their ballot manifests and CVRs
- RLA Tool will compile uploaded county data in separate database tables on back end

Implementing RLAs (continued)

- RLA Tool will “ping” the back-end ballot manifest and CVR tables and randomly select ballots to audit
- County audit boards will retrieve appropriate ballots and report the votes using the RLA Tool
- If the reported votes match the corresponding CVR, the ballot “passes” and the county proceeds to the next randomly selected ballot
- If the reported votes do not match the corresponding CVR, the ballot “fails” and additional ballots may need to be randomly selected to satisfy risk limit

RLA Challenges & Shortcomings

- Physical duplication of damaged/electronic ballots
- Digital adjudication of ballots with ambiguous markings in accordance with the Secretary of State's Voter Intent Guide
- Maintaining accurate ballot manifests
- Maintaining tabulated ballots in order they are scanned
- Very limited time (5-9 days) to start & complete RLA
- Detailed CVRs versus ballot secrecy and voter anonymity
- Colorado's RLA is a tabulation audit rather than an audit of the entire election. For example, RLA does not audit or verify signature verification of accepted mail ballots

Conclusion

- This is going to be an iterative process
- We expect to de-brief with our consulting subject matter experts and county partners to collect lessons learned, modify procedures going forward, etc.
- Questions?