Risk-Limiting Audits (RLA)

Neal McBurnett, Harvie Branscomb
For the Free and Equal Electoral Reform Symposium
2019-12-07

Overview of the Journey

- Post-Election Audits are Important
- Evidence-Based Elections and BMDs
- RLA vs other kinds of Audits
- Data Format Standards
- How RLA Works in CO – The Basics
- Status of RLA Process in Colorado and Beyond
- Using RLA with Alternative Voting Methods
- Web Resources

Why Audits are Important

- Ensure that votes are counted accurately and securely, while protecting voter privacy. Want to **confirm** election **outcomes** and **correct** errors.
- Machine interpretation is recorded in a Cast Vote Record, but machines can fail and also might misinterpret ballots marked by humans.
- Routine audit in Palm Beach County, FL in 2012 revealed two city council contests were certified with the wrong outcomes. See also RI, PA, etc.

Evidence-Based Elections

- We can't certify the accuracy of computer-based voting systems
- We must audit and certify each contest in each election => **Evidence-Based Elections**
- **Risk-Limiting Audits (RLAs)** check the **tabulation** of a set of voter-verifiable paper ballots
- RLAs should be used in conjunction with other audits (**ballot reconciliation, signature verification, chain-of-custody**, etc.) to support Evidence-Based Elections

Auditable Voting Systems

- Top priority is having **voter-verified paper ballots**
- Traditional **Hand-Marked Paper Ballots (HMPB)** are in best conditions already verified
- **Ballot Marking Devices (BMD)** were introduced for accessibility, hard to get voters to verify paper
- The easiest tabulation method to audit is **central count** of HMPB (e.g. Colorado)
  - Possible to check the match of each anonymous ballot with a corresponding Cast Vote Record
- More common are **precinct-count** which require much more auditing

BMDs and Voter Verification

- Ballot Marking Devices (BMDs) produce paper ballots that voters often don't actually look at or verify.
- Vulnerable to presentation attacks, printing a ballot which doesn't match what was shown on screen
- The less voters verify, and the more ballots are cast via BMD, the weaker the overall evidence and the more the actual outcomes remain vulnerable to hacking
- Stationing an election worker in front of each scanner, encouraging voters to look at and verify their paper ballot, is one way to increase verification rates
- => **Minimize BMD use, or innovate on verification**

Kinds of Tabulation Audits

- Fixed Percentage – Example: 2% of precincts
- Fixed Size – Example: 1,000 ballots
- Tiered Samples
    - depending on reported margin of victory
- Risk-Limiting Audits
- End-to-end open audits via ElectionGuard. cf. STAR-Vote, Scantegrity

Why Risk-Limiting Audits are Better

- We want vote counts to be at least accurate enough to produce a corrected outcome
- Traditional tabulation audits usually either
  - require more work than necessary to confirm an outcome, or
  - yield too little information to be conclusive.
  - An RLA uses statistics to check voted ballots until it has strong evidence that election outcome is correct according to the evidence provided to them. Then the audit can stop. Efficient!
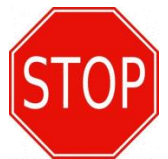
Definitions:
Types of Risk-Limiting Audits

- **Ballot comparison** – audit individual ballots
  - Verify that the Cast Vote Record (ballot specific machine interpretation) is correct
- **Batch Comparison** – audit entire batches or precincts (less efficient and less informative)
- **Ballot Polling** – random sample of ballots if auditable counts aren't available. Less efficient by factor of 1/margin

Challenges

Why is it taking so long to adopt robust audits?

- Elections are increasingly complicated
- You can't easily audit the data you've got
- You can't easily get the data you need
- Critical *Common Data Standards* work by Election Assistance Commission (EAC) / NIST
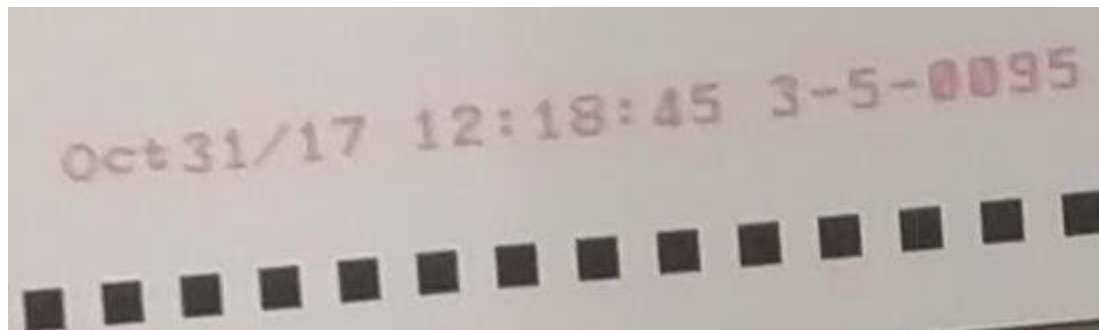
Source: http://bcn.boulder.co.us/~neal/elections/rla-nasem.pdf

A Ballot Selected for Audit

A Ballot and its tracking sheet
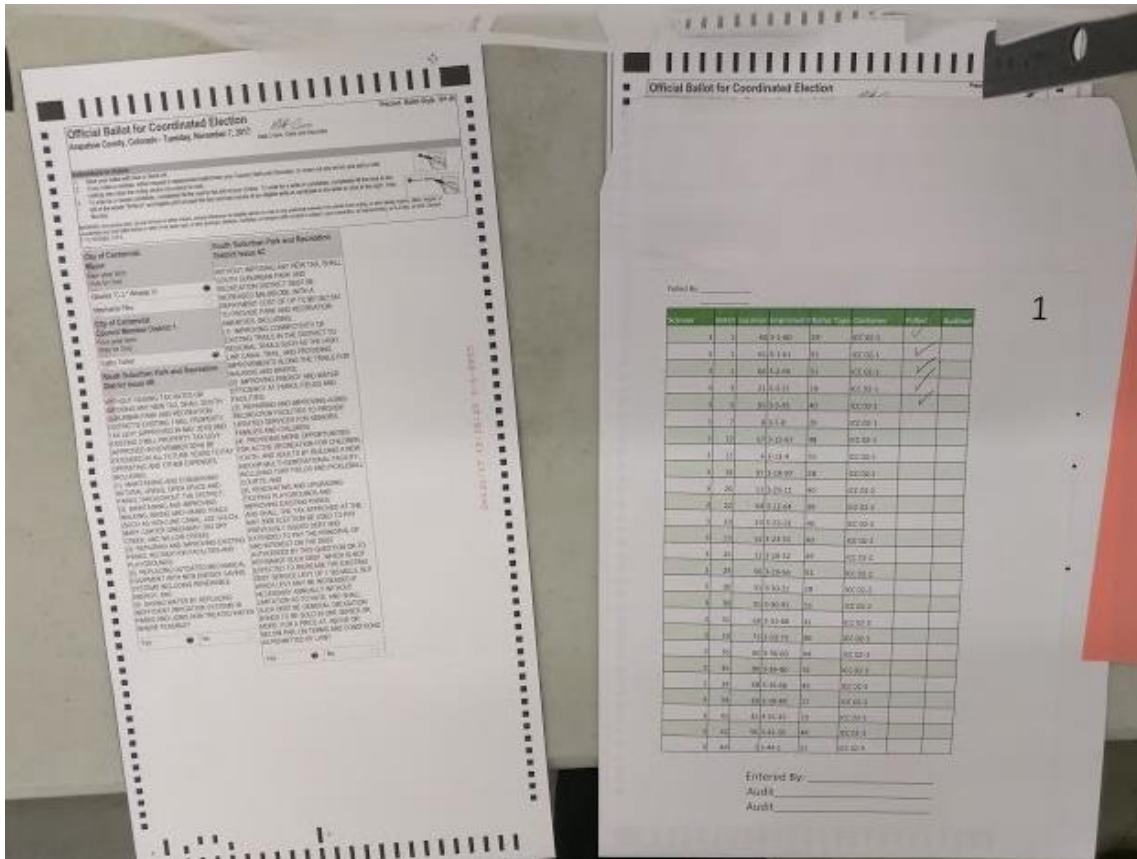
Ballot Identification

Imprinted ID showing a ballot was scanned on October 31st at 12:18:45 and positively identified as ballot card "**3-5-0095**": scanner 3, batch 5, 95th card

Data Format Standards: Critical!

You can't easily audit the data you've got

- Need to look thru all 200,000 ballots to find the ones you selected
- Different formats, often undocumented or "proprietary", previously from 4 different vendors in CO

You can't easily get the data you need

- Generate a full report for each batch, may need to calculate batch totals from differences
- Implemented in Boulder, starting in 2008 election
- Big step forward with first open source code (ElectionAudits) and Boulder County Clerk Hillary Hall and her amazing team who made it their own in following years (Excel etc)

Source: http://bcn.boulder.co.us/~neal/elections/rla-nasem.pdf

Common Data Formats

- We need format standards! OASIS. IEEE. EAC/NIST
  - John Wack: Overview of VVSG-Interoperability Common Data Formats (two presentations)
- Election Results CDF V1 published as SP 1500-100.
  - Used in OH, NC, LA County, other states in progress.
- V2 synchronizes with Google/VIP 5.1, adds JSON.
- Election Log Export CDF soon published as SP 1500-101.
- Voter Records Interchange CDF slated for review by VR vendors and then published as SP 1500-102.
  - Initial use in OH and by OSET.
- Cast Vote Records CDF schema approved by WG, to be published as SP 1500-103.
- Continued development and documentation of election process business models and voting method descriptions.

Source: http://bcn.boulder.co.us/~neal/elections/rla-nasem.pdf

Evidence presented and checked

- Detailed Public RLA Oversight Protocol, Stephanie Singer, Neal McBurnett 2017

- Elements:

1 Chain of Custody
2 Tabulation
3 Manifest
4 Commitment
5 Random selection
6 Ballot card retrieval
7 Ballot Interpretation and data entry
8 Ending the random selection and examination of ballot cards
9 Hand Count
10 Audit Conclusions Affect Outcomes

See http://bcn.boulder.co.us/~neal/elections/PublicRLAOversightProtocol.pdf

Public engagement in verification

- Promote public participation in audit
  - e.g. attend dice roll and manual interpretation
  - Access to scanned images and CVRs

- Print ballot tracking pages with QR codes

- App to photograph ballot + QR code

- Assist public tweets like "I verified this vote"

Example of a misinterpretation

Early Summary

* In audits of "driving" contests: 20 possible discrepancies out of 3015 audited ballot sheets

* Of 4 investigated discrepancies, 3 seem like entry errors.
  Note entry is blind, and no feedback is given at the time.
  4th was inconsequential: a vote that shouldn't be counted for sole candidate in contest

DIRECTOR DISTRICT 2
Four Year Term
(vote for one)

Susan Harmon

● Erica Shields

DIRECTOR DISTRICT 5
Four Year Term
(vote for one)

Ron Mitchell

Definitions:
Risk-Limiting Audit Theory

- **Risk Limit** – largest statistical probability that an incorrect reported tabulation outcome is not detected and corrected in a risk-limiting audit. Worst-case scenario! E.g. 5%, 20%
- **Diluted Margin** – the smallest margin (in any contest) as a fraction of all the ballots subject to the audit
- Vote **Overstatement** (narrows the margin) and Vote **Understatement** (increases the margin)
  –Based on pairwise margins in a contest
  –Over or under by 1 or 2

Definitions: Logistics

- **Publicly Verifiable Random Seed** – a starting point for randomly selecting ballots to audit
  - A 20-digit number, e.g. 84437724778708423271
  - 20 stakeholders each roll a 10-sided die.
  - Put the 20-digit number into a public pseudo-random number generator to determine which ballots to audit from a prepared ballot manifest

Public Meeting to Establish the Random Seed – Nov 9, 2017

Video: https://youtu.be/SU8kYvsQCC0
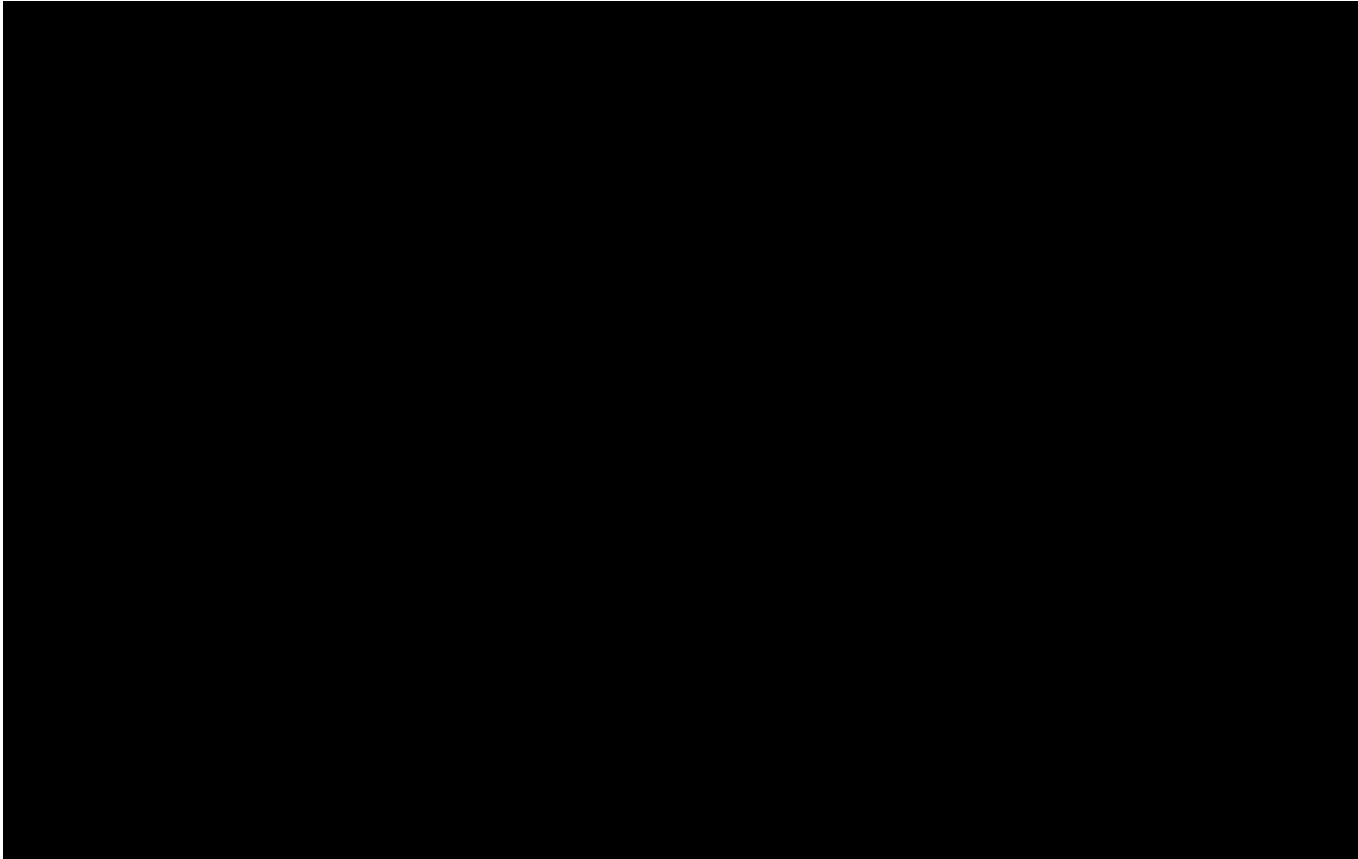
Definitions: Logistics (continued)

- **Ballot Manifest** – a list detailing where each ballot is located

Ballot Manifest (Excerpt)
Boulder County

| County	Device ID	Batch	# of Ballot	Location
BOULDER	1	1	146	1
BOULDER	1	2	142	1
BOULDER	1	3	147	1
BOULDER	1	4	140	1
BOULDER	1	5	142	1
BOULDER	1	6	139	1
BOULDER	1	7	147	1
BOULDER	1	8	147	1
BOULDER	1	9	133	11
BOULDER	1	10	141	11
BOULDER	1	11	144	11
BOULDER	1	12	146	11
BOULDER	1	13	146	11
BOULDER	1	14	144	11
BOULDER	1	15	149	11
BOULDER	1	16	145	11
BOULDER	1	17	150	21
BOULDER	1	18	149	21
BOULDER	1	19	119	21

Definitions: Logistics(continued)

•**Ballot Cards** – individual pieces of paper that together constitute a single ballot containing all of the contests an elector is eligible to vote

`

How RLA Works in CO – The Basics
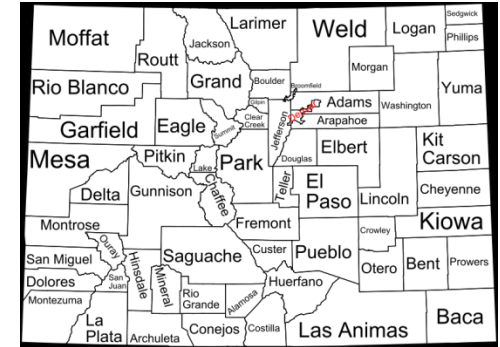
- Breakdown in 2019:
  - 62 counties: Ballot Comparison
  - 2 counties: Hand Count Ballots
- Targeted only 1 contest per county, 1 statewide
- Others could be audited "opportunistically".

Status of RLA in CO: Successes

- Efficiently-auditable tabulation
- All contests subject to audit (but not reviewed)
- Open Source Software developed for ballot-level RLAs
- Publicly verifiable random selection
- Officials could check risk measurements

Status of RLA in CO: Remaining work

- End SOS (especially as a candidate) from responsibility to choose audited contests
- Target the most interesting, closest contests, to a larger risk limit if necessary, and target all remaining contests with a reasonable estimated workload
- Share results for opportunistic audits. Allow public RLA oversight (publish CVRs, `rla_export` data, images)
    - Requires addressing anonymity issues better
- Handle non-voter-verifiable ballots properly (e.g., received by email)

RLAs in Other States

- Auditing more challenging in most states: in-person scanners require randomization of CVRs => can't match with paper ballots
- Require **Batch Comparison** or **Ballot-Polling audits**
- New **Arlo** software, in Python
- Math for more efficient Ballot-Polling audits

RLA Laws, Pilots Spreading

- RLAs now required by law in CO, RI, NM, CA (for some equipment)
- Pilots done in IN, VA, MI, NJ, RI, VA, PA, OH, GA and more

Using RLA with IRV or STV

- In instant-runoff voting (IRV) or single transferable vote (STV), even determining the margin (minimum number of changed ballots that could lead to different outcome) is tricky!
- Groundbreaking 2019 IRV audit, San Francisco
- Bayes audits are more flexible, provide metrics for any voting method, but not always risk limits
- No traditional frequentist approach is available for STV yet.

RLA and
Single-Winner Voting Methods

- Conventional (easy)
- Approval (easy)
- Score (SHANGRLA)
- STAR Voting (SHANGRLA)
- Cumulative Voting (SHANGRLA)
- Instant-Runoff Voting (SHANGRLA+RAIRE)

RLA and
Multi-Winner Voting Methods

- Conventional tabulation (easy)
- D'Hondt and related Proportional Representation (SHANGRLA)
- Sequential Proportional Approval Voting (Bayes)
- Score (SHANGRLA)
- Single Transferable Vote (STV) (Bayes)
- STV with reduced runoff (Bayes)

Auditable Voting Systems

- Top priority is having **voter-verified paper ballots**
- Traditional **Hand-Marked Paper Ballots (HMPB)** are in best conditions already verified
- **Ballot Marking Devices (BMD)** were introduced for accessibility, hard to get voters to verify paper
- The easiest tabulation method to audit is **central count** of HMPB (e.g. Colorado)
  - Possible to check the match of each anonymous ballot with a corresponding Cast Vote Record
- More common are **precinct-count** which require much more auditing

Website Resources

- CO Risk-Limiting Audit Project (CORLA):
http://bcn.boulder.co.us/~neal/elections/corla/

- CO Secretary of State Audit Center:
http://www.sos.state.co.us/pubs/elections/auditCenter.html

- *A Gentle Introduction to Risk-Limiting Audits*
https://www.stat.berkeley.edu/~stark/Preprints/gentle12.pdf

Website Resources

- Software for auditing:
https://github.com/ron-rivest/ElectionAuditWareRepo

- Harvie Branscomb's Election Quality
website: http://electionquality.com/

- This presentation:
http://bcn.boulder.co.us/~neal/talks/rla2019.pdf