Steps in the conduct of a risk limiting vote tabulation audit only for counties with publishable CVRs for CO SOS

draft H Feb. 12, 2017

Working assumptions: (see Glossary on page 3 to understand acronyms)

All ballots are on paper
Ballot format differs- sometimes in-person is different from flat or mail format
All ballots are scanned (and have mCVRs)
Ballots not able to be scanned are duplicated and then scanned
Ballots are transported to central count location
Ballots are in collections determined by jurisdiction conducting election
Some ballots will arrive late (UOCAVA, cure, wrong county drop)
There are serious time constraints between last tabulation and canvass deadline
Ranked voting methods not supported; plurality, threshold and approval supported
Ballot interpretation is guided by "voter intent guide" (specified by SoS)
Some ballots are not anonymous

- For reasons of rare ballot style or format
- For reasons of voter created identifiable marks

Goals:

-- risk limiting audit of every contest that is county-wide to pre-specified risk limit
-- for the other contests, an opportunistic risk-measuring audit
-- observability of audit process to public (including candidates)
-- efficiency of audit; accurate results for minimum workload under available constraints
-- triggers full manual recount after reaching a threshold of escalation
-- all audit steps and decisions generate public records, actions visible by public

Simplified steps: County perspective

1. Publicly announces time and place for audit;
2. works with SOS to predict and resolve any privacy issues related to rare ballot styles;
3. learns from Audit Center (AC) how many ballots to plan to audit, how many to hold back;
4. reports ballot manifest to Audit Center daily- style, sequence, batch, box, room, county;
5. Election night -reports ENR (election night report of tallies) and mCVRs to Audit Center;
6. finds updated estimate of number of ballots to audit on Audit Center for workload planning;
7. reports final ENR and final mCVRs to Audit Center upon final post election tabulation;
8. accepts proposals for targeted audit units, vets, then creates targeted aCVRs as convenient;
9. finds ballot sequence to sample on Audit Center plus look ahead to later samples;
10. locates, opens, captures voter intent on paper and creates aCVR for required samples;
11. if duplicated, find original and create separate aoCVR for it;
12. uploads manually captured voter intent from sampled paper ballots (aCVRs,aoCVRs) to AC;
13. checks Audit Center for any additional samples needed based on discrepancies, targeted audits, and margins of larger contests;
14. return to step 9 until Audit Contest Manager agrees audit is settled or in recount; and
15. reports process, resources, workload and exceptions to Audit Center.

Simplified Steps: SOS perspective

1. Publicly announce time and place for dice rolling to generate random seed;
2. receives contest info and assesses privacy and need for initial samples;
3. publishes predicted initial sample sizes and numbers of ballots to hold back for late tabulation for each county;
4. republishes ballot manifests received from counties, daily, preferably by script;
5. receives ENR and mCVRs from counties election night, calculates margins;
6. estimates sample size needed from each county, publishes estimates;
7. receives, republishes updates on chain-of-custody, ballot manifest, mCVRs daily until final;
8. conducts public dice roll to determine PRNG seed, map of counties to statewide range;
9. publishes required list of samples by manifest entry per county including future samples list;
10. checks uploaded audit aCVRs for comparison to machine mCVRs, determines need for escalation;
11. coordinates with Contest Audit Manager to decide about recount or satisfaction of contest audit;
12. publishes remaining active contests, requirement for additional samples per county, if any;
13. loops on step 9 until all contests are in recount or audit settled; and
14. publishes report of audit with risk measurements for all contests.

Glossary:

- Active contest: a contest for which the desired risk level has not yet been reached. Uncontested contests are never active.
- Audit Center: a SOS run web-presented information center for statewide coordination of audits
- Ballot manifest: a list of ballots with a descriptor to associate each with its style and precise physical location (e.g., box number(s)) [see example ballot manifest at [URL]]
- Ballot position: numerical descriptor used to locate ballot in a batch (e.g. count from bottom)
- Ballot style: a descriptor that characterizes only the set of contests on a ballot
- BBB (new term): Ballot/Batch/Box identifier – tells where the ballot can be found- may include ballot position, batch name, box name, room and county
- CVR: Cast Vote Record - documented pattern of captured voter intent that is capable of being aggregated
- aCVR (new term): Audit Cast Vote Record (manually created record of voter intent for auditing)
- aoCVR (new term): Audit Originally Cast Vote Record - the ACVR for original duplicated voter intent
- mCVR (new term): Machine Cast Vote Record (machine + human adjudicated CVR in digital form)
- DEO: Designated Election Official- the authority over coordination of contests on a ballot
- ED: Election Day – the statutory deadline for most ballots to be returned in Colorado
- ENR: Election night report of tallies by counties to the SOS
- IRV/RCV: – Instant Runoff Voting / Rank Choice Voting and other Ranked Voting Methods are not supported by this audit
- Mandatory contest - contest for which risk limit will be required to be met
- Opportunistic contest - contest for which the risk will be measured and not necessarily met
- Presumed loser: or runner-up; in unaudited or unofficial results, the contest choice that is not reported as winning but would be the first to win if election tallies changed in a way unfavorable to the reported winner
- PRNG: Pseudo-Random Number Generator – open source software to produce pseudo random integers
- RCV: see IRV; rank voting methods are not supported by this audit
- RLA: risk limiting audit – an audit that limits the risk that an incorrect outcome will not be corrected
- RLTA, RLVTA: risk limiting (vote) tabulation audit – an RLA of presumed eligible ballots
- Settled contest or audit: contest for which the measured risk is below the limit and need not continue
- SOS: Secretary of State

Roles:

- Audit Board
- Candidates, Initiative Committee Chairs
- Canvass Board
- County Audit Manager
- County Staff
- DEO - Designated Election Official, County Clerk and Recorder
- Multi County Contest Audit Manager
- Secretary of State
- SOS Staff
- Statewide Audit Manager

Reports:

	"C" means county originated, "S" means SOS originated "*" means an existing report	Sample Links
C	**Pre-coordination Electors per Style** and Format	
C*	**Contests on Ballots** (final ballot design)	
S	**Ballots to Withhold** per county per style for privacy	
S	**Planned Initial Sample** size per county	
C	**Daily Ballot Manifest** (ballot position/ batch number/ box number/ room/ county)	
C*	**Seal Logs Report** for storage of ballots, envelopes, stubs, EMS backups, ballot scans, etc.	
S	**Ballots On Hand**- daily report of per county by style (copied from daily ballot manifests)	
C*	**ENR Upload** – tallies of tabulated ballots reported to state	
C	**mCVR Upload** - machine provided cast vote records for entire county (possible privacy substitutions)	
S	**Election Night Ballot Manifest** statewide with "privacy concern" entries flagged	
S	**Election Night mCVR** report- substituting for privacy concern entries	
S	**Preliminary Audit Sample** size by county based on election night results	
S	**Random Seed and Sequence** - random generated integers to be used for sampling statewide	
S	**Order Of Counties** and # ballots for mapping to pseudo-random generated integers	
S	**Audit Samples Requested** from ballot manifest in order of selection by county	
C	**aCVR** manual cast vote records from audited ballots; including original intent, revised and targeted	
S	**Contest Tallies** made by aggregating mCVRs (excluding privacy concern mCVRs)	
S	**Discrepancy and Contest Settled Report** (issued each round of auditing)	
S	**Final Statewide Audit Report**	

Steps begin here:(days of week are assuming a Tuesday election day)

These steps assume that each county treats its ballots as a single collection that is closed and committed to and evidence (mCVRs) published before a single run of a PRNG. An alternate approach is to have more than one collection per county, and more than run of a PRNG. Such a split approach would allow audit samples to be chosen before late ballots have returned, but audit coordination would be more complex.

Pre-election

1. Build a SOS-run centralized web information center (Audit Center or "AC") that can be used by 64 counties and various municipalities and other districts who opt-in to perform the RLTA. Audit Center specifies the format for daily upload of ballot manifests as ballots are tabulated (starting 15 days prior to ED). Audit Center specifies the manner in which original evidence of voter intent is to be kept so that it can be easily accessed in case a duplicated ballot is selected for audit. Audit Center specifies the format for upload of CVR's and for coordination of contest names to match ENR (election night reporting) so that CVR reports can be aggregated across counties including any last-minute changes due to litigation, etc. Audit Center supplies aCVR forms in meant-for-printing and electronic versions for election judges to use to capture voter intent from paper.
2. The Audit Center and counties publicly announce the times and places for the random draws and the audits so that they can be observed by the public. The source for the PRNG to be used is published.
3. Collect names/identifiers for all contests to be placed on county coordinated ballots plus any standalone district ballots and agree on standard metadata to be used by all participants for the RLA, including contest, candidate, and data field names, and permissible values for each field.
4. Each county prior to making a decision regarding coordination of a non-precinct based district determines if there are instances where less than 30 ballots of a single style will as a result be issued to voters. For each such case the relevant counties and DEO shall have a conference with the SOS officials representing the Audit Center to determine if there is a way to avoid issuing less than 30 of a specific style. One such method would be to put one or more contests on a separate page that can be given a separate style not to be correlated with the precinct based page and not to be combined into a single mCVR. Such prior planning will assist the efficiency of the audit because fewer or even zero ballots might need to be withdrawn from the audit and fewer ballots will need to be audited as a result.
5. Based on number of contests and estimated number of voters for each contest, Audit Center assigns an initial sample size for each county and participating standalone district – to be published on the SOS Audit Center site.
6. Audit center reports publicly to counties how many election night mCVR/ballot combinations per each style are to be withheld for privacy reasons until 9th day for late counting. (Envelopes to be checked for eligibility and opened but ballots not scanned or otherwise tabulated.)
7. Counties plan to hire election judges and or canvass board members to attend post-election process that will start on 10th day and continue on the 13th with an expectation to conduct double the initial sample.
8. Audit Center recommends to each county to remove identifiable marks from paper ballots through duplication or removable occlusion prior to scanning. This will also make the audit more transparent since scans of these ballots can be released to the public. For counties with rare use of the QR format, selections-only ballot, Audit Center recommends that these ballots be duplicated onto flat mail-in type ballots so that their rarity will not cause them to be kept confidential for reasons of voter privacy.

Tabulation starts

9. County documents the location of each tabulated ballot in each batch in each box on the BBB ballot manifest form and uploads to Audit Center daily. These ballot/batch/box (BBB) reports also are accompanied by copies of seal logs for chain of custody of ballots, envelopes, stubs, backed up electronic

data such as images and mCVRs as well as any data substantiating the eligibility of the ballots collected (e.g. reports of eligibility audit, etc.) Each individual BBB entry informs how to find a specific ballot on hand.

10. Audit Center publicly reports the count of sheets of tabulated ballots per county by style with a daily update

Election night, 7PM

11. Counties begin upload of conventional ENR reports to SOS and to Audit Center if separate. Shortly thereafter, counties begin to upload mCVR files from the voting system to match the ballot/batch/box BBB manifest reports already uploaded and still being uploaded. Counties withhold from tabulation sufficient batch collections to satisfy the quota for late counting as specified earlier by Audit Center in step 6.
12. Before end of business election night, counties if practical upload all matching CVR files for previously uploaded ballot/batch/box manifests. Ballot/batch/box manifests are immediately published by Audit Center. Individual CVRs are not published at this time. County informs AC when all ballots on hand have been tabulated and CVRs uploaded.

Election Day plus 1 (Wednesday)

13. Counties finish uploading mCVRs. Audit Center runs a software tool to count ballot styles from mCVRs and creates a list of ballot styles and "potentially identifiable" mCVRs so-named because under 10 examples exist among uploaded mCVRs. (10 is the current statutory threshold for anonymity)
14. Audit Center replaces "potentially identifiable" and otherwise withheld mCVRs with obvious substitute entries (e.g. all overvotes) but each in keeping with the correct style of the original CVR. Audit Center publishes all CVRs including those substituted for reason of protection of privacy on the SOS public web site- each CVR should still match to a BBB manifest entry.
15. At some point after ED a mechanism for requesting audit capture of targeted ballots opens to the public. To avoid blind target requests, this process will depend on public access to paper and or scanned images in addition to CVRs. Implementation of this access may take some statutory adjustment and may be a year or more off in the future. The original scanned image data or BBB manifest entry for the paper ballot and mCVR data used for the comparison by the public must be included in the request for targeting transmitted to the Audit Center. This request is input by the public into a web form that collects the necessary data such as the BBB entry for each ballot identified as mis-adjudicated as well as the mCVR and the AC-published scanned image. If this publicly uploaded data does not match the digital signature of data published by the Audit Center, then the request for inclusion in the audit as a target is disregarded. For a ballot to be added to the targeted sample, a suggestion is that it must have been identified by two separate entities as discrepant to avoid spurious requests. County audit managers may schedule voter-intent-capture of targeted paper ballots at their convenience prior to completion of the audit, and this process may begin after election day and also prior to sampling decisions by the AC for the required portion of the RLTA. This process for targeted audit can be implemented in a future election if appropriate.

Election Day plus 2 (Thursday)

16. Audit Center calculates margins for all contests based on mCVRs by replacing substituted mCVRs with votes for the presumed losing choice. Audit Center then determines based on available mCVRs the narrowest margin contest countywide in each county.

Election Day plus 3 (Friday)

17. Audit Center calculates the predicted actual initial sample size for each county based on narrowest margin countywide contest. This decision isn't final until late ballots are tabulated but it will help plan how many judges will be required to process audited ballots.
18. Counties continue uploading BBB manifest reports as late ballots arrive. Counties report seal log changes for any existing BBB reports to document chain-of-custody. In some years this is Veterans Day holiday. In otheryears it is ED + 6 a Monday.

Election Day plus 8 (Wednesday)

19. Deadline for arrival of late ballots (UOCAVA, signature verification and missing signature and missing ID cures, plus ballot packets dropped in boxes out of county).

Election Day plus 9 (Thursday)

20. Counties mix, batch and tabulate the remaining ballots together to maintain privacy protection. The last of the mCVR reports are uploaded to the state referring to the appropriate BBB entries that these ballots belong with. BBB manifest entries of late counted ballots are not organized (batched) by origin or format of ballot. Counties each confirm final upload of CVRs to the AC. Tardy counties to be contacted immediately by state support personnel to provide assistance.
21. Audit Center replaces with correct values the substituted mCVRs for all but the "<10-styles" withheld ballots and republishes all the mCVRs such that contests choices can be easily summed to produce election tallies. Also, if possible, Audit Center (if not already) after calculating digital signatures for each collection of scanned images, publishes the anonymous scanned images except for those that match the "<10-styles" mCVRs (and any other paper/scan/CVR that has been withheld for privacy reasons). Audit Center recalculates the minimum county-wide margin contests assuming the remaining non-published ballot/CVRs are voted against the presumed winner. The Audit Center publishes new initial sample requirements for each county (minimum being the originally published sample requirements from Step 5).
22. For each county, the final full county set of mCVRs is converted into a calculated "fingerprint" (hash) equivalent to 20 decimal digits.
23. A pre-announced public ceremony is held at the SOS office and streamed live. The numerical seed to be used for random selection is created by attending stakeholders and public each of whom throws a 10-sided die or equivalent. The equivalent of 20 throws are required to produce a 20 digit decimal value. This is appended to the value of the digital signatures calculated in Step 22. The resulting seed is input during the ceremony by AC staff into an open source and well-specified randomizer (published and known-good PRNG) that calculates values from 1 to the total number of ballots in the statewide BBB manifest. The resulting list of integer values are printed out and published on the Audit Center site. This list will be used to select samples from the concatenated final county BBB manifest reports. The sampling of the audit is completely reproducible and verifiable if the same seed is used with the specified PRNG and if the mapping of counties' order in the list remains the same. This is a design feature that provides reproducibility. An alternate scheme involves the AC running the PRNG 64 times each time with a range of values appropriate to the length of each county manifest list. In this case for each run of the PRNG the digital signature value calculated for the county in Step 22 is appended to the result of the statewide 20 digit dice roll.
24. If the alternate 64 county list scheme is not used, in the same public ceremony, dice are thrown again to determine the order of placement of counties' BBB manifest lists in a master statewide manifest list. This is the audit equivalent of cutting the deck after shuffling. Not needed for technical criteria of sufficient randomness but a visible source of uncertainty that any observer can recognize.
25. AC associates the ballot manifest BBB list with the county portion of a PRNG list such that PRNG values index into the county list and publishes the resulting list of paper ballots to be audited for each county, each list long enough to accommodate all expected audit needs of the county short of a complete manual recount.

Election Day plus 10 (Friday)

26. Initial samples - paper ballots - are pulled at each county, election judges capture voter intent for each active contest from each ballot, and "audit" cast vote records are manually created (aCVRs). Possible captured entries are "vote for X", "undervoted", "overvoted", "damaged", "exception". If the selected tabulated mCVR and ballot were subjected to duplication or re-making, then the original evidence for the voter intent is accessed and a separate aoCVR is created (audit originally Cast Vote Record). aCVRs and aoCVRs can be marked on prepared paper forms or entered into the county audit management software or into spreadsheets where each sample ballot is given a row and each contest is a column on a spreadsheet or several 11x17 landscape printed pages (the data for these pages can be downloaded from Audit Center.)

Counties may go ahead and sample additional ballots from further down the sample list if they have time and sufficient resources.

27. Counties upload a(o)CVR sheets (by scanning the paper or transmitting the data) to the Audit Center where these are published. Once the aCVR is uploaded, the equivalent sheets representing the mCVRs (machine Cast Vote Records) for the matching BBB manifest entries are published by Audit Center. The format of these sheets are such that they can be optionally printed and, if desired, visually compared by superimposing the sheets on a light table or in front of a bright light to detect discrepancies. Other means of comparison may be used including machine-assisted comparison if manual verification is also possible.
28. The Audit Center publishes a report by county including all aCVRs and aoCVRs in the election such that comparison to corresponding mCVRs is possible. Counties may begin investigation of discrepancies - all of which should have been reported initially by AC and any special cases reported by counties to AC. If counties have reason to, a revised audit cast vote record may be uploaded to the AC, always accompanied by a statement of the reason for the revision. Discrepancies are resolved through the AC reporting process.
29. The mCVR report for all contests and all CVRs from step 21 is tabulated and compared to the ENR reports for final election tallies. (Note that these reports will differ by the aggregate contents of the CVRs removed for privacy protection. This is a technical defect with the audit that would best be prevented by designing the ballot design and collection such that it does not require data suppression to protect privacy.)
30. aCVR/mCVR pairs are compared and discrepancies are evaluated for statewide and multi county contests and county wide contests and the risk limit for each contest is measured. Risk limits that have achieved or superseded the pre-established limits for each contest stop the audit for those contests, rendering them "settled."
31. For remaining contests that need to be further audited, the discrepancies in individual CVR contents between aCVR and mCVR found by Audit Center are evaluated and used to call for additional samples in order to reach risk limits in county-wide and larger contests.The list of active contests for each county is updated by AC to remove those for which the audit is "settled." Additional samples are then moved from the projected sample list to the required list for counties for which risk limits are not met.

Election Day plus 11,12 (Weekend)

32. These days will be used by counties if needed for escalated auditing, particularly for a primary election.

Election Day plus 13 (Monday)

33. Selected counties for which additional samples are required bring in election judges and paper or spreadsheets are used to collect voter intent from the selected remaining contests for which audit data is still needed. The goal of this day's work is to complete any necessary auditing above the county level. If once the day's samples are uploaded and compared to mCVRs and further escalation is needed, a state-level audit management meeting is held to determine if one or more recounts should be initiated. For a primary election today is the last day for canvass. This might mean that for some contests the audit remains "active" or in recount.
34. At the end of this day the Audit Center publishes the measured risk for all intra-county contests.
35. Counties who wish to do so may escalate their own audits to achieve desired higher risk limits or to satisfy local campaigns with respect to measured risk deemed inadequate for an intra-county contest. All such additional auditing and the aCVRs associated are uploaded to the Audit Center.
36. SOS creates a report of the measured risk for all contests. Counties report the basic process used for creating aCVRs and for locating the sampled ballots and any exceptional conditions encountered. SOS publishes a comprehensive report of the tabulation audit.
37. The deadline for canvass in a general election is ED + 17; in a primary election ED + 13; ED+16 is always Thanksgiving for a November election. ED+15 is the more realistic deadline. ED+3 or ED+6 is Veterans Day holiday in a November election.

Policy decisions to be made (incomplete list):

Architecture

1. How to handle counties without mCVR but with batch or precinct or device tabulation?
2. How to handle hand count counties?
3. What are the pre-set risk limits and for what categories of election (statewide, countywide, legislative)?
4. Are standalone district elections supported by the RLVTA?
5. How to handle contests smaller than countywide?
6. When, if ever, to update election results based on discovered discrepancies?
7. What triggers transition from audit escalation to recount?
8. When review voter intent guide based on differences of opinion on voter intent reflected in reported discrepancies?

Mechanics

9. Sampling decisions made only centrally or in addition in each county?
10. Single run of PRNG statewide or 64 runs of PRNG one for each county?
11. Single or multiple collections of ballots per county (each collection independently subject to sampling)?
12. Decide to sample a collection before all scanning complete? Interpret ballots before late ballot returns?
13. Trust PRNG output to be unaffected by insiders? (Expect external confirmation of PRNG list?)
14. How account for "privacy concern" ballots in sampling and upon selection?
15. How will "privacy concern" ballots, if selected for audit, be treated to protect privacy?
16. Is ballot manifest generated at opening or by the eligibility check system, or?
17. How and where will the authoritative comparison of aCVR to mCVR be done?
18. Enforce blind re-interpretation for audit?
19. Prevent re-interpretation of same ballot by same election judges for audit?

Targeted Ballots and Public Access

20. How and when to inform public about ballot content to inform the targeting of ballots for audit?
21. How decide to accept targeted ballots/contests for extra audit steps?
22. When begin to audit targeted units?
23. Audit only targeted ballots before late ballots come in?

Software (and major functionality) Required

SOS-run centralized web information center (Audit Center or "AC")
Provides information and regulations about the statewide risk-limiting vote tabulation audit (RLVTA) and a web interface from which participating counties, municipalities and districts can download forms (e.g., ballot manifest, optional aCVR form meant for printing) and data (e.g., lists of specific ballots to be sampled, standard contest names, etc.) and which those authorized administrative jurisdictions can also use to upload data (e.g., contest and candidate names, mCVR's, aCVR's, etc.).

This software should be mirrored on different hardware with different URL's so authorized administrative jurisdictions have guaranteed fast, secure access while the general public can access a separate copy. There also need to be alternative ways to accomplish the necessary tasks in case the web information center should become unavailable at any time.

Pseudo Random Number Generation (PRNG)
Given an initial numerical "seed", this software provides lists of pseudo random integer numbers that can be used by the AC web interface to specify ballots to be sampled. Several free, open-source PRNG packages are available, including the "sampler" program from Ron Rivest, also used by Philip Stark's online software.

Other software required at the state level
Software to restructure and merge uploaded mCVRs so that each multi county district election can be tallied as a sum of all component county mCVRs.
Software to support sampling decisions (how many to sample based on margins, etc.) comparable to Philip Stark's online software.

Software required at the county level, if any, beyond software already used for tabulation, ballot management, etc.
Spreadsheets or paper or custom software may be used to capture voter intent from paper ballots for uploading to AC. If county based sampling is used, PRNG and sampling decision software will be needed.

Major Contributors (without implication of endorsement)

- Harvie Branscomb
- Neal McBurnett
- Ron Rivest
- Lynn Garland
- John McCarthy