Check All Election Contests, with Several Options

November 6, 2022

Audits or recounts of all election contests, on all ballots, are important for public acceptance of election results. The public, or trusted representatives of each group which may doubt results, need to see enough steps to catch and correct problems. Neither insiders nor computers are usually trusted.

Accidental errors are probably more common than hacks, and cannot be deterred. The only solution is to check as thoroughly as budget allows. If hackers know an election will be checked, they are less likely to hack it, so checking provides benefits even if nothing is found.[1]

Checking everything can be done at a reasonable cost. It would be a real improvement for states like Colorado, Massachusetts and Washington, which only do full audits of a few contests, or like California and Texas, which exclude many votes from their audits..[2]

Checking all contests recognizes that millions of dollars in contracts and land use decisions are at stake in every contest. We don't want criminal groups picking local prosecutors, sheriffs or judges. We don't want crooks or extremists picking school boards or city or county councils. They set policies for police, schools, housing codes and other key aspects of daily life. They decide on millions of dollars in contracts and affect more millions of dollars in property values.

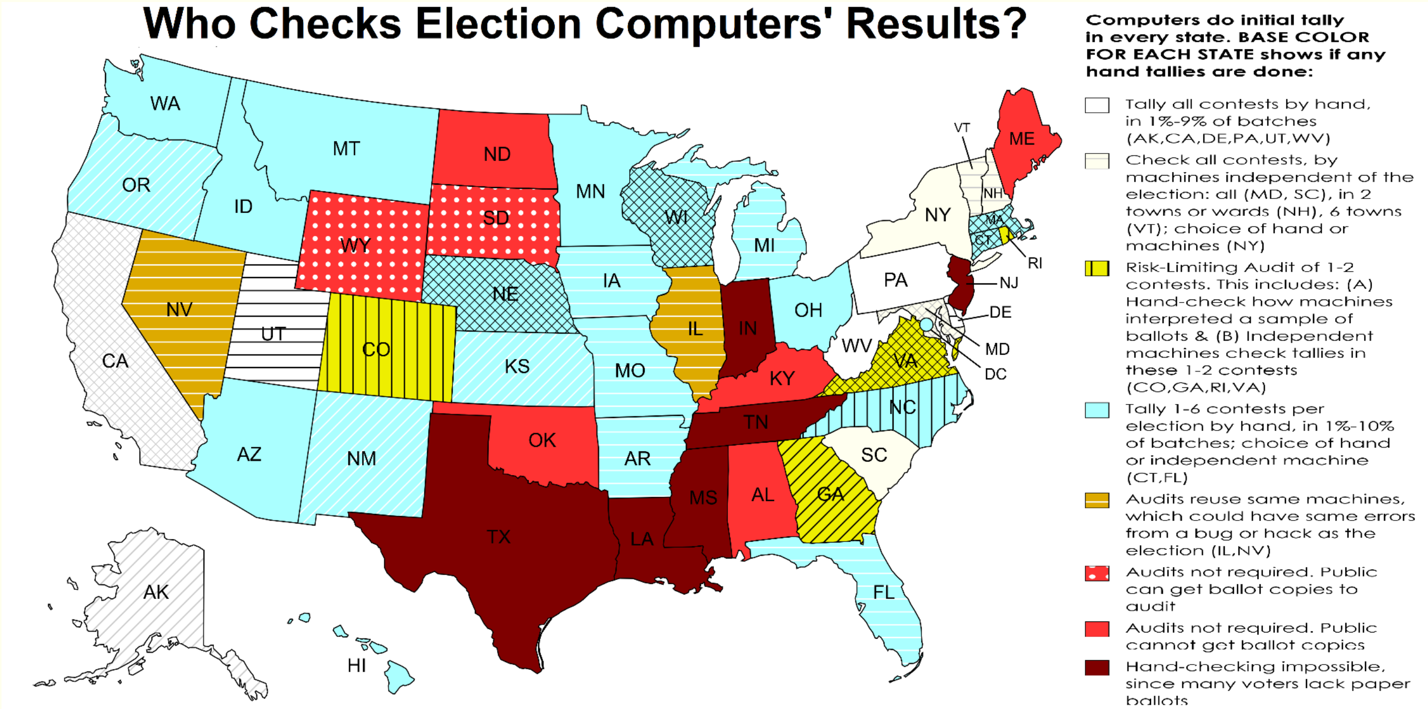
Even small samples are more reliable than not checking contests at all.

[1] The approaches in this paper are described more briefly and graphically in Count Votes; Check, and in an animated video.

[2] State audit reports and rules are at http://www.votewell.net/audits.html . PA's traditional sample is like CO, DE, IL, IN, NV, NY, UT, WV, which audit all contests on a sample of all ballots, so they cover a sample of local contests. PA's settlement with Jill Stein requires "audit all unofficial election results using robust pre-certification audit methods..." https://storage.courtlistener.com/recap/gov.uscourts.paed.524624/gov.uscourts.paed.524624.108.1_1.pdf . However PA has decided to stay with its traditional sample of batches with 2% or 2,000 ballots, whichever is less in each county. PA calls it an RLA, but it has no risk limit and no expansion if the initial sample does not confirm the outcome. https://www.vote.pa.gov/About-Elections/Pages/Post-Election-Audits.aspx

STATE RULES FOR AUDITS, BALLOT IMAGES & CONFIDENTIALITY

State rules on audits



All states tally votes with computers, which have bugs, and can be hacked by insiders or annual updates. If ballots or copies are made available, independent tallies can check results.

DIAGONALS mean some types of ballots are excluded from checking: AK,OR-small precincts; CA,KS-ballots not tallied by election day; CT,WI-early+absentee+provisional ballots; GA,MA,MD,NM,VA,WI-primaries; GA,MA,NE-off years

HORIZONTAL lines: laws do not require results to be corrected when audits find errors: AR,CA,CT,DE,FL,IA,IL,MA,MD,MI,MO,NE,NH,NV,TN,UT,VA,VT,WI

CROSSHATCH: both the above problems: CA,CT,MA,MD,NE,VA,WI. Note: MD & PA have another audit, besides what is shown: MD hand checks one state-wide contest; PA is starting RLAs.

VERTICAL lines: audited contests have good samples of ballots, likely to catch any significant errors by machines interpreting ballots in the races checked: CO,MD,NC,NM,PA,RI,VA

Sources: State laws and rules for 2020-22. Cited, with color map, at commons.wikimedia.org/wiki/File:State_audits.png

State reports and rules on **audit** methods, samples, and timing are at http://Votewell.net/audits.html

State rules on ballot images

Local practices on **preserving** ballot images are at https://commons.wikimedia.org/wiki/File:Keepimages.png

State rules on public **availability** of ballot images: https:/commons.wikimedia.org/wiki/File:Ballot-foia.png

Public access while protecting vote secrecy & confidentiality

Comparison of approaches

Model laws for selected topics on audits, confidentiality, accessibility.

The first option to provide independent review of election counts is to release complete unredacted raw files directly copied from voting machines, with no processing or re-formatting, with digital signatures, to court under seal, and/or to pre-vetted organizations with penalties for disclosure. There needs to be a priority decision whether this happens before or after processing and publication of preliminary results. In either case, the release can be within hours after polls close, to minimize time for loss or mistakes. Lawyers, CPAs, people with security clearances and many academics have experience keeping files confidential. Other options, which are not mutually exclusive, are to post these files for the public, or redact distinctive marks, handwriting, and rare ballot styles and then post for the public. San Francisco takes a month to redact, and doing a good job probably needs to take that long, which makes the files too late for correcting official results.

When ballot images or cast vote records (CVRs) are available to the public, there are many options to address confidentiality of each voter's votes. Many experts on many subjects have developed approaches to let outsiders analyze data and still protect confidentiality of individual records. Options include dividing the information so no one can link any of it to a person (such as putting images from each contest into a separate file, to tabulate separately), nondisclosure agreements, redaction (which is sometimes required already before scanning), or full release with prohibitions on some uses.

Some observers say that public files let voters sell their vote by proving how they voted. They could identify their ballot image by marks, handwriting or a special name written in, or an unlikely pattern of votes. If people want to sell votes, they can do so with mailed ballots, far more effectively, by handing the ballot to the buyer to submit at the last minute, when it cannot be changed. Large vote buying is hard to keep secret. Separating images or CVRs from each contest into a separate file removes the risk of identifying one vote by anything marked on another contest. It is hard to buy a vote for Jan Doe for Sheriff by any handwriting or entry in the write-in space, since that would no longer be a vote for Sheriff Doe.

RISK LIMITING AUDITS (RLAs) ARE AN OPTION

RLAs to audit all contests

RLAs have up to now declined to audit all contests in any election.[3] They could do so, and costs would be reasonable.

As an example, the Allegheny County, PA 2019 general election had many town & school elections. Table R1 below shows the sample size needed for RLAs to audit all town elections with 10% risk limit.[4] Towns don't overlap, so the town samples are independent and can be added. Similarly school districts don't overlap and their samples are added in table R1. Towns and school districts do overlap. The sampling rate in each area needs to be the higher of sample needed to audit the school election there and sample needed to audit the municipal election there. Naturally some school elections which need high sampling rates are in areas where town elections need small sampling rates, and vice versa. The overall sample needed is therefore bigger than either estimate, and less than their total. The estimates of cost and sample as a percent of voters are based on adding half the school district sample to the municipal sample, to have adequate sample in each contest.

The sample needed in a jurisdiction does not depend on the average sample in the many wide-margin contests, but on the closest margin, hence biggest sample, in each local area, which is then enough for all other contests in that area too. This summation of samples needed in non-overlapping areas has only been done in these elections in Table R1.

The Allegheny County, PA 2020 general election had state legislative elections. The table shows RLA samples needed to audit all state representatives and both state senators in the county. The overall sample needed is therefore bigger than either estimate, and less than their total.

---

[3] Colorado transcribes all contests on the ballot sheets it samples, and reports the number of discrepancies for individual contests. It does not limit risk nor draw conclusions about outcomes, except for two contests on each ballot. Samples are typically small, so they omit contests which are not on those ballot sheets.

[4] Risk limit is 10% in (many are at https://www.electionsecuritywa.org/learn-more):

Stark, 2010 Super-Simple Simultaneous Single-Ballot Risk-Limiting Audits

Lindeman, Stark, Yates, 2012 BRAVO: Ballot-polling Risk-limiting Audits to Verify Outcomes

Lindeman, Stark, 2012 A Gentle Introduction to Risk-limiting Audits

Risk-Limiting Audits Working Group, 2012 Risk-Limiting Post-Election Audits: Why and How

Stark, Wagner, 2012 Evidence-Based Elections

Schein, Brown, 2021 Risk-Limiting Audits: A Guide for Global Use. "risk limits generally seem to be 10 percent or less" https://www.ifes.org/sites/default/files/ifes_risk-limiting_audits_a_guide_for_global_use_march_2021.pdf

State Audit Working Group, 2021 Comments to Senate http://votewell.net/S1ReconciliationComments202107.pdf

Risk limit is 5% in:

Stark, 2018 An Introduction to Risk-Limiting Audits and Evidence-Based Elections Prepared for the Little Hoover Commission

California law https://www.sos.ca.gov/administration/regulations/current-regulations/elections/risk-limiting-audits

Other:

"Colorado began RLAs with a 9% risk-limit in comparison audits, and has reduced that number to 4% as counties got more comfortable performing the audit." https://www.sos.state.co.us/pubs/elections/RLA/faqs.html

Rhode Island "estimated work loads for both 5% and 10% risk limits" 2019 p.46 https://www.commoncause.org/rhode-island/wp-content/uploads/sites/26/2019/08/RI-Report-Design-FINAL-WEB5.pdf

Lack of public trust. A 2021 survey of 105 US residents[5] found they did not trust the small samples[6] provided by RLAs in most contests. They wanted a median 6.9% sampling rate. Over 40% of respondents wanted at least a 10% sampling rate. If one purpose of audits is to reassure voters, RLA samples are too small. Even the 7% RLA sample calculated for the Allegheny 2019 general election does not give 6.9% in most contests. A few close contests would have had 100% samples, with most under 3%. When respondents were told the RLA sample sizes, and were asked if such an audit would increase their confidence in the election result, 45% said "maybe no" or "definitely not." A few respondents volunteered, without being asked, that they were concerned about the reliability of the paper trail. The respondents were more educated than average; 66% had a college degree. The study did not examine voter confidence in other approaches, such as retallying 100% of images, and checking that a sample of images are accurate, or 100% rescanning and reanalysis (Options described later in this paper).

The Hennepin County, MN, 2021 general election had town and school elections. The table shows RLA samples needed to audit all municipal elections (mutually exclusive) and school elections (mutually exclusive, though overlapping with towns). As with Allegheny County, the overall Hennepin County RLA sample needed is bigger than either estimate, and less than their total.

Minneapolis, St. Paul, Bloomington, Minnetonka, and St. Louis Park used ranked choice voting (RCV). The numbers here only handle the winner among first choices. Additional rounds may or may not need bigger samples, which are not generally possible to determine from published data.

Table R1. RLA sample size	Staff cost in this county	Election	Total samples, 10% risk level	Sort key in Column AN of votewell.net/tallies.xlsx	Ballots	RLA sample as % of ballots	Number of contests where 1% is enough for RLA
14,782	$78,000	PA_Alleg_2019g	Municipal contests	m1	276,484	7.0%	9 of 72
9,337		PA_Alleg_2019g	School contests	s1	276,484		12 of 30
693	$27,000	PA_Alleg_2020g	State Reps.	LR	726,720	0.1%	14 of 15
466		PA_Alleg_2020g	State Senators	LS	726,720		2 of 2
2,042	$23,000	MN_Henne_2021g	Municipal contests	m1	180,000	1.8%	13 of 21
2,209		MN_Henne_2021g	School contests	s1	180,000		6 of 11

Source: "costs" and "tallies" tabs in http://votewell.net/tallies.xlsx and "MN,PA" tab in costs spreadsheet
RLA sample sizes would actually be bigger, by the (unreported) proportion of overvotes & undervotes.
RLA estimates above exclude city-wide contests when a ward needed a bigger sample. The citywide race would still need extra samples in the other wards.
Estimates exclude any extra samples needed to audit additional rounds of ranked-choice voting in MN.
Samples would be drawn from cast vote records per arxiv.org/abs/2012.03371 . Estimates exclude phantom votes described in that paper.
Formula for ballot comparison RLAs: (6.2 / margin_as_fraction_of_ballots_in_contest), from p.44 of www.stat.berkeley.edu/~stark/Preprints/RLAwhitepaper12.pdf More efficient designs have not been codified for nonspecialists

The calculations in both states depend on sampling cast vote records.[7] That method requires first counting poll book signatures, mail ballot approvals and provisional ballot approvals. These counts need to be done separately for each ballot style, to find the maximum number of ballot sheets voted in each contest. If this maximum is more than the CVR file shows, auditors create phantom CVRs for the ballot sheets missing from the CVRs. The audit treats phantoms as votes for the loser. Finding the maximum for each ballot style will be hard to get right, for temporary workers in the paper blizzard of elections and the rush of post-election processing.

Voters who walk out after being issued a ballot, without casting it, would create this excess of ballot sheets issued, beyond the CVRs, plus miscounts at any step. And even a few phantoms can put a close contest into question. If phantoms exceed the winning margin in a contest, or otherwise put it into question, auditors need to manually examine all ballot sheets in the jurisdiction to ensure they find all sheets which have the contest. The estimates here assume no phantom CVRs, and would be bigger when records suggest phantom CVRs are needed.

After checking samples, ballot comparison RLAs have a software step, to independently sum computer interpretations of ballots (cast vote records, CVRs), to check the election system's totals.[8]

---

[5] Dalela, Kulyk, Schurman. "Voter Perceptions of Trust in Risk-Limiting Audits" https://arxiv.org/pdf/2109.07918

[6] Doubts about RLA sample sizes in actual practice have also been voiced by Mercuri https://www.cnet.com/news/electronic-voting-and-partial-audits/

[7] "More Style, Less Work" 2021. https://arxiv.org/abs/2012.03371 (p.15 discusses phantoms.)

[8] The basic requirement in RLAs to check vote totals by some means is in:

Need for sample sizes

The large RLA samples are driven by a few close contests in each election. If you believe there could be just enough hacked ballot images to create these close local wins, and few or no hacked images elsewhere (where the rest of the jurisdiction's or state's sample would find them), then you need something like the RLA sample sizes shown. Otherwise smaller samples, such as 1%, are enough to see if hacked images have moved out of the lab into real elections.

RLA advocates have not in fact generally called for auditing any of these close local contests at all.

The Colorado League of Women Voters has endorsed "Evidence-based election auditing for all contests, including some risk-limiting audits."[9] The last column of table R1 shows that checking 1% of all ballots would comply with their goal by providing a large enough sample for "some risk-limiting audits" in each election.

Table R2. RLA sample size, examples	Ballots in contest	Winning margin, votes	Place
Municipal races, mayor or council, 2019 General Election, Allegheny County, PA			
2,296	2,296	1	Jefferson Hills (vote for 3)
1,086	5,604	32	Plum (vote for 3)
946	2,747	18	Franklin Park ward 2 (vote for 1)
School directors, 2019 General Election, Allegheny County, PA			
3,408	3,408	4	Hampton (vote for 5)
2,015	2,015	4	Riverview (vote for 5)
1,176	11,379	60	North Allegheny (vote for 5)
1,085	5,776	33	Plum (vote for 1)
Municipal races, mayor or council, 2021 General Election, Hennepin County, MN			
1,133	9,501	52	Minneapolis-Council Member Ward 2
School board races, 2021 General Election, Hennepin County, MN			
1,308	9,492	45	ISD #276-School Board Member
Source: "MN,PA" tab in costs spreadsheet			

Checking stored ballots

RLAs are not designed to find flaws in the stored ballots they use. They call for other procedures to check storage.

It is *not* always possible to see directly whether skilled crooks bypassed locks, seals, security cameras and chain of custody records to change stored paper ballots.

We *can* detect illicit changes by having accurate electronic copies of ballots before the ballots go into storage and using these verified electronic files to check if paper ballots were changed during storage. The verified electronic files ensure that fold lines and scanner lines on ballot images do not create false votes, ensure more accurate initial results, and let us recover true copies of ballots if needed.

Quality control while scanning: Ballots can be, and usually are, safely under bipartisan observation, which prevents changes, before the ballots are scanned. Places which store mailed ballots unscanned, out of sight of observers and staff, lose this benefit of continuous bipartisan observation unless they have bipartisan guards 24/7. They will have no way to check whether ballots are changed in storage.

Lindeman+Stark, Gentle Introduction to Risk-limiting Audits. p.3, section B(i) "*ballots sum to the contest totals for every candidate.*" https://www.stat.berkeley.edu/~stark/Preprints/gentle12.pdf.

Item 4b *"reconciliation to ensure that all votes from all audit units are correctly summed in the election totals*", on page 12 of https://electionaudits.org/files/Audit%20Principles%20and%20Best%20Practices%202018.pdf

Lindeman et al. Comments re statistics of auditing the 2018 Colorado elections. p.2. "*the Secretary of State has the ability to compare reported contest results with the CVRs it receives, but there is no publicly observable means to verify that this comparison has been done correctly.*" sos.state.co.us/pubs/elections/VotingSystems/riskAuditFiles/NotesOnStatistics-2018ColoradoPrimariesElectionAudit.pdf

[9] LWV of CO, Election Security Position. 2022. https://www.lwvcolorado.org/docs.ashx?id=940583 and https://lwvcolorado.org/content.aspx?page_id=22&club_id=314195&module_id=506542

Election scanners create an electronic record of each ballot. At time of scanning, before ballots go into storage, staff and observers can compare a random quality control sample of ballots to the newly created electronic records, and correct problems. These steps ensure the electronic records are true copies of the paper ballots. These electronic records are what the election system uses to tabulate results, so checking while scanning is a best management practice, regardless of whether audits are done.

Quality control software checks *all* scans to see if they have unreadable barcodes, stretched images, spurious lines (black, white, from folds, bad sensors, etc.) or other issues which software can detect.

Besides 100% quality checks by software, quality control also includes humans comparing a sample of paper ballots to electronic records to detect other issues while the paper ballots are still at hand, before they go into storage.

Detailed steps while scanning are described in the section, "QUALITY CONTROL FOR SCANNERS."

Each defect at this quality control stage needs to be resolved, until we are confident the electronic files match the paper ballots. This stage ensures initial election results are as accurate as possible. These accurate files provide a baseline to detect any changes during storage. Copies of the verified electronic files go into separate safes, of multiple elected officials or judges, to cut the chance an accident or hack could harm all copies at once.

Checking RLA samples after storage: For RLAs, auditors go into storage rooms to bring out a sample of stored ballots. Auditors also need to get the electronic files corresponding to these ballots and run the same computer algorithm which calculated the original hash values. If a hash value does not match, auditors need to get a copy of the file from another safe and be sure its hash value matches the original.

Auditors audit the sample ballots they pulled from storage. They also need to compare these ballots to the electronic files. If any stored ballots do not match the verified electronic files, auditors need to investigate further, and may need to replace changed ballots with better copies from the electronic files.

Combined sampling & storage risks

Most audits have sampling risk. For example an RLA may have 95% chance of catching inaccuracies in tabulation. The original file verification (at time of scanning) also had sampling risk. For example it may have had 90% chance of catching an error level big enough to change a particular contest's outcome. In this situation, the overall chance that the audit using stored ballots will catch inaccuracies is 95% times 90%, or 86%.

BALLOT IMAGE AUDITS (BIAs) ARE AN OPTION

Modern election machines create an electronic image of each ballot.[10] This Option checks if these ballot images match paper ballots, and also uses the images to check elections, without risk to the original paper ballots.

1. Workers check if ballot images are accurate, by comparing a random sample of ballot images to paper ballots. If any ballots have errors in the images, workers investigate and correct the causes.

2. Software, independent of the election system, re-interprets and retallies *all* contests on *all* ballot images, to check if every vote was interpreted correctly, and checks if all tallies are right (by batch, contest and district), even for small, close contests.

3. Humans examine all ballots where the BIA software and election software disagreed, to see which is right and to improve both systems for the future. If overvotes, write-ins and faint marks are enough to flip an outcome, humans review them too.

Even if step 1 is not completed, a ballot image audit of this type addresses most risks.

[10] Equipment from Dominion, Hart, Clear Ballot and ES&S high-speed scanners save all images, as a general rule. ES&S precinct scanners have settings to save all images, none, or just write-ins.

How to check if ballot images match paper ballots

Ballot images need to be checked. Like all computer products, electronic ballot images created by election scanners can have errors, accidental or intentional. Some image errors are visible on the image, such as dropping out areas of the ballot, and white[11] or black[12] streaks from faulty sensors. Software or humans can find these issues without even looking at the paper. Staff and observers compare a random sample of paper ballots to ballot images to find other issues, which have not been seen in actual elections, but could happen, such as missing faint marks, or shifting votes between candidates.[13] It is hard to justify examining large samples of paper ballots for not-yet-seen problems, but jurisdictions should do as much as they can afford. Sufficient checking deters hackers and limits accidental errors.

Table E1 below shows how to check a sample of paper ballots to find and correct flawed images. It's important to report all flaws found. This reporting motivates other jurisdictions to audit (many audit nothing[14]), encourages them to find the best scanners and provide the regular maintenance needed, and prompts them to prosecute wrong-doing if some flaws were not accidents.

[11] Gideon, John (2005). "Hart InterCivic Optical-Scan Has A Weak Spot". *www.votersunite.org*

[12] Walker, Natasha (2017). "2016 Post-Election Audits in Maryland" (PDF). *Elections Advisory Commission*.

[13] Image errors were found in a compression system used in some Xerox scanners up to 2015. Kriesel 2015. "Images are cut into small segments, which are grouped by similarity. For every group only a representative segment is saved that gets reused instead of other group members, which may cause character substitution. Different to PM&S, SPM corrects such errors by additionally saving difference images containing the differences of the reused symbols in comparison to the original image. This correction step seems to have been left out by Xerox." (SPM means Soft Pattern Matching which is needed for a lossless version of the JBIG2 image format used by Xerox.) https://www.dkriesel.com/en/blog/2013/0802_xerox-workcentres_are_switching_written_numbers_when_scanning

Image hacking has been done in the lab with a truncated ballot and a Windows scanner driver: "UnclearBallot: Automated Ballot Image Manipulation." 2019.
https://www.researchgate.net/profile/David-Duenas-Cid/publication/336000609_Electronic_Voting_4th_International_Joint_Conference_E-Vote-ID_2019_Bregenz_Austria_October_1-4_2019_Proceedings_4th_International_Joint_Conference_E-Vote-ID_2019_Bregenz_Austria_October_1-4_2019_Proc

[14] https://en.wikipedia.org/wiki/Election_audit#Table_of_**audit_rules** and https://verifiedvoting.org/auditlaws/

	Table E1. Basic Approaches to Check if Paper Ballots Match Electronic Records		
	Ballots scanned centrally (E.g., mailed ballots)	**Ballots scanned in polling places** (4 other approaches for these ballots are in Table E2 in the section “QUALITY CONTROL FOR SCANNERS”)	**Jurisdictions without ballot images** (another approach is in Table E2 in the section “QUALITY CONTROL FOR SCANNERS”)
	(a)	(b)	(g)
1	When each batch is scanned, take the steps below to examine a 1% sample of ballot sheets.[A]	Rescan all polling place ballots centrally. Then treat like other centrally scanned ballots.	Require that images be saved wherever the scanner can do so. Then use the approaches on the left. Otherwise, take steps below to check the files of Cast Vote Records (CVRs). For (g) rescan polling place ballots centrally.
2	Central office publishes hash values[B] of image files (and/or CVRs, as available). These are provided on paper to anyone present and online in a time-stamped archive, to deter and detect any later alterations to the files. Hash values and sealed electronic files are put in multiple safes to cut the chance that all are hacked.		
3	Throw three 10-sided dice (marked 0-9, 00-90, and 000-900), to select the sample of ballots.[A]		
4	Count to those random numbers in stack of paper ballot sheets fresh out of the central scanner.		
5	Count in the electronic file to the corresponding ballot images or Cast Vote Records (CVRs).		Compare sample of ballots to their Cast Vote Records (CVRs).
	If any part of the image or CVR differs from the paper ballot, (a) report differences, (b) fix the scanner or use another,[D] (c) rescan, and (d) re-post the hash. It is important to throw the dice again to get a new sample to test the corrected file.[E] Using a new sample avoids checking the same ballots, so no one can suspect that staff just corrected the few ballots in the sample. Staff can repeat these steps until images are accurate.		If any vote differs between the paper ballot and CVR, (a) report differences, (b) investigate to see if they can be fixed (poor scanner or software creating CVRs), (c) fix problem, (d) re-process, (e) re-post hash, (f) re-throw dice, until CVRs are accurate; or (g) hand-tally all ballots.
7	Investigate any differences to see if they affected preceding and later ballots, find causes, and find preventive measures.		
8	**Independent software.** After election day and when most ballot images are available, (a) run independent software, (b) compare its interpretation of the votes on each image to the election system’s interpretation, (c) report differences, and (d) tally totals to check outcomes in every contest.		
	Table Notes [A] **Sample size** needs to be chosen in each state or jurisdiction. It is a trade-off between cost and accuracy, as discussed in the section “Sample sizes.” In small counties (or in towns administering their own elections) the county’s (town’s) sample should have at least 100 sheets. [B] A **hash value** is a fixed-length string of numbers and letters generated from a mathematical algorithm and is unique to the file being hashed, as defined in CISA ST04-018. https://www.cisa.gov/tips/st04-018 [D] If a **discrepancy** happened because the scanned image lost faint marks, may need to change to gray-scale, color, or higher resolution. [E] **After finding one problem,** the office may want to sample more than 1% on re-do.		

How to check all contests on all ballots

Independent software. After polls close and most ballot images are available, independent software can examine the images and tally results to check the election system and individual contest outcomes. Starting early lets problems be found and resolved as soon as possible. Ballots arriving later can be added when the final ones arrive, to complete the analysis.

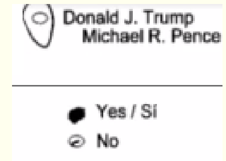
In a similar way, bank checks are scanned immediately, and scans are trusted from then on. The paper checks are shredded, though ballots would not be.[15]

[15] Checks are shredded: "check truncation" https://www.ffiec.gov/exam/check21/faq.htm

"The lack of access to original checks necessitates the need for revised check review procedures and employee training, and a review of check security features such as watermarks." https://www.fdic.gov/news/news/inactivefinancial/2004/fil5404.html

The state or each jurisdiction can have an open competition for independent software or develop the expertise at a local institution. Images have been independently interpreted by Clear Ballot in CO and MD,[16] by UCSD and Election Transparency Project in CA, by AuditEngine in FL, GA and WI,[17] and by Audit Station in CT.[18]

The most trustworthy approach is for any candidate or party to hire its own independent organization to examine and tally images and bring forward any evidence they find of discrepancies, as described in the Option for recounts.



Poor software can miss circle and count hesitation dots.

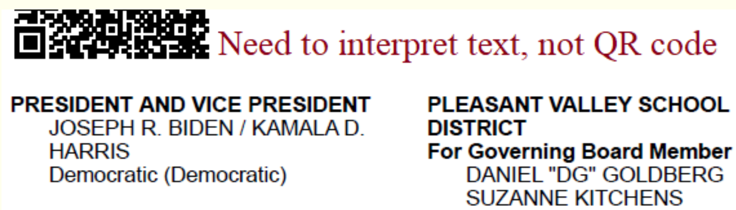
How humans check discrepancies

Independent ballot image analysis software is just as subject to errors and hacking as other election software, but generally would not make the same errors. Therefore it is best to examine and document all discrepancies, image by image. Then judge which software is correct for that instance and why the other software made its mistake, so both software systems can be improved.

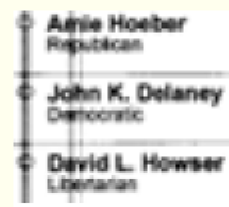
When evaluating multiple contests that have discrepancies, the most urgent to resolve are the contests where the number of discrepancies is enough to affect an outcome. Resolving discrepancies in other contests helps improve software before problems are big enough to change outcomes.

The independent analysis needs to handle common problems. Several common problems involve overvotes. When there are enough overvotes to change an outcome in any contest, and if state law requires following the intent of the voter, then overvotes need manual review.

Need to interpret text, not QR code



- A circle, checkmark, X, etc. can show the voter's intent even if it does not darken the target area. The software needs to analyze outside the target area and process marks near it or display them for human review.
- A mark, where another mark is crossed out, does not need to be treated as an overvote.
- A mark for a write-in space does not need to be treated as an overvote, if the name written in is also marked as a listed candidate, or did not register as a write-in candidate (in states which require it), or no name is written in. State law may provide rules.
- A small mark, such as a dot, is not a vote if there are darker marks on the same contest.
- Software can tally voters' marks and the text, not bar codes or QR codes[19] which are on some ballots.



Flawed scanner put black lines on image, making overvotes

Image problems. Besides discrepancies in tallying votes, the software also needs to look for image problems, such as wide black or white streaks which cover voting locations. If the problems are common, the random sample finds and corrects them. The 100% examination of ballot images finds the problems even if they are rare, which protects results in small close contests. These kinds of problems will not necessarily appear as discrepancies, since both official and independent software may interpret black lines as overvotes and white lines as undervotes. The software needs to look for streaks, fold lines and dropouts specifically. If the software

---

[16] Clear Ballot independently interprets Maryland's ballot images. There is no direct check that images match paper ballots. MD counties check one statewide contest to see if net results of paper ballots match CVRs in 5% of precincts. This check can miss image errors on other contests and offsetting image errors. Checking all contests on some paper ballots would give a better check than 1 contest on 5%.

Clear Ballot also independently interpreted El Paso County, CO, 2020 ballot images. El Paso compared all contests to CVRs on 1,195 ballots and found 25 discrepancies. They did not report if images or software caused these discrepancies, but in general Colorado trusts the images (p.18).

[17] Cost and availability of independent software to tally electronic ballot images is at http://www.votewell.net/auditcosts.html

[18] Antonyan, Tigran; Bromley, Theodore; Michel, Laurent; Russell, Alexander; Shvartsman, Alexander; Stark, Suzanne (2013-06-21). "Computer Assisted Post Election Audits", *State Certification Testing of Voting Systems National Conference*.

[19] Ballot marking devices (BMDs) let voters enter choices on a screen or with buttons, to print a ballot which shows the votes and often a bar code or QR code summary. Some voters with disabilities and others use BMDs.

detects this kind of problem and the lines are so wide they completely hide some votes, staff can go back to the paper ballots, use the rest of the image to verify the paper ballot has not been altered in storage,[20] and determine the voter's intent.

Hand tallies. Ballot images can be hand-tallied, like paper ballots, if someone wishes to, and without any risk to the original paper ballots.[21]

Recounts. Whether or not the election office hires an independent company to tally all contests on all ballot images, it could allow candidates to access the image files and to recount for themselves (see Option for recounts). In large contests this is much cheaper for the candidates and less work for the staff than recounting with paper ballots.

PROS & CONS OF RLAs & BIAs

First, RLAs base their sampling rates on margins of victory in each contest, which means waiting for final results and **using stored ballots**. RLAs call for checking the chain of custody, but they have not proposed a reliable way to test for surreptitious access. They usually call for checking logs, which are not necessarily stored securely and which cannot reflect surreptitious access.[22] Image-checking can generally examine paper ballots much earlier, when each batch of early ballots is scanned. Thus the paper ballots are examined when they are reliable, before they are subject to the alterations and losses while ballots are in storage, which have happened regularly throughout our history. Sampling can start this soon, because it uses an overall sampling rate, chosen and budgeted for, before the election starts.

RLAs can, though they don't yet, add a step of verifying electronic files before paper ballots go into storage, then checking during the audit that ballots still match the verified electronic files.

Current RLAs accept the risk of using paper ballots, which could have been altered in storage, a risk with a long, large history in US elections,[23] while they avoid the risk of image hacks.

---

[20] There are no statistics on how often criminals enter rooms undetected, but law enforcement often does so, so ability to enter rooms undetected is widespread at least in law enforcement and former law enforcement: Electronic Frontier Foundation, "Peekaboo, I See You: Government Authority Intended for Terrorism is Used for Other Purposes". Also McGuire, Sneak and Peek Warrants-Necessary for our Safety...?

[21] Hand-tally of all contests would use teams of two reading to teams of two talliers, not sort-and-stack, which is less accurate: Goggin, Stephen N.; Byrne, Michael D.; Gilbert, Juan E. (March 2012). "Post-Election Auditing: Effects of Procedure and Ballot Type on Manual Counting Accuracy, Efficiency, and Auditor Satisfaction and Confidence". *Election Law Journal: Rules, Politics, and Policy*. **11** (1): 36–51. doi:10.1089/elj.2010.0098. ISSN 1533-1296.

[22] Stored ballots have parallel weaknesses to ballot marking devices (BMDs). Appel, DeMillo and Stark note, "the essential security flaw of BMDs: few voters will notice and promptly report discrepancies between what they saw on the screen and what is on the BMD printout, and even when they do notice, there's nothing appropriate that can be done. Even if election officials are convinced that BMDs malfunctioned, *there is no way to determine who really won*." ("Ballot-Marking Devices (BMDs) Cannot Assure the Will of the Voters." *SSRN Electronic Journal* 2019.)

The parallel issue is that: No voters will notice if their paper ballot is changed in storage, and there's nothing appropriate that can be done. Even if election officials are convinced that locks, seals and security cameras were breached, *there is no way to determine who really won*.

Appel previously wrote, "I demonstrated for the judge the complete removal and replacement of all seals [on voting machines] with no visible evidence of tampering..." Security Seals on Voting Machines: A Case Study . *ACM Trans. Inf. Syst. Secur.* 2011. Johnston & Warner have said, "no seal is unspoofable (just as no lock is undefeatable)." "How to Choose and Use Seals" *Army Sustainment.* 2012. In 2020 Appel was in a committee which said, "RLAs require paper ballots or records, and a degree of chain-of-custody over ballots that few states and local jurisdictions currently require." https://www.law.uci.edu/faculty/full-time/hasen/2020ElectionReport.pdf

Stark was co-author on a paper, Bernhard, Benaloh, Halderman, Rivest, Ryan, Stark, Teague, Vora, and Wallach. "No state has laws or regulations to ensure that the paper trail is conserved adequately, and that evidence to that effect is provided." https://mbernhard.com/papers/voting-sok17.pdf

Halderman was co-author on a paper about creating key blanks from a photo of a lock: Burgess, Wustrow & Halderman; 2015. "Replication Prohibited: Attacking Restricted Keyways with 3D-Printing"

[23] **Examples of unauthorized access to ballots and other election records:** http://www.votewell.net/locks.html

The alternative of picking a percentage in advance to check as ballots are scanned, avoids the risk of using paper ballots which could have been altered in storage, and accepts the risk of missing small, focused image hacks, which have no history in US elections. State and national clearinghouses are important to hear if image hacks do appear. Most governments allocate most money to avoid known risks, and small amounts to watch for new risks.

Second, a result of RLA post-election timing is that an RLA catches errors after the initial erroneous results are already released. This is embarrassing for officials and mistrusted by candidates who lose votes. Checking while scanning corrects errors before results are released.

Third, RLAs use samples to find all errors, so all RLA findings have a **risk level**. With images, most errors are checked by the independent software looking at all contests on all ballots, so there is no sampling error on these findings. There are only sampling errors in the detection of hacked images, which are still only theoretical.

Fourth, both approaches are **hard to observe**, in different ways. RLAs require staff to go into storage rooms, unobserved by the public, to bring back boxes of paper ballots. BIAs do not. Then RLA staff check ballots by entering them into software, and observers have generally not been able to see the ballots for themselves. With images, observers would be able to see the projected ballots during checking, only if they attend multiple scanning sessions over weeks. Even in an RLA, these scanning sessions are where the chain of custody is created, and they need to be observed, which is time-consuming. Observers are rare at any stage. So in both cases success depends on honest bipartisan election staff and their ability to report issues, as discussed below under HOTLINES.

Fifth, when an RLA cannot confirm official results, it calls for a **100% hand count**, which is slow, expensive, and hard to manage. Maryland decided not to use risk-limiting audits, because of the high sample size on close contests and chance of 100% hand tallies.[24] When image checking finds image problems, it calls for rescanning on a better scanner or a repaired scanner, which is easier than a complete hand count. When image checking finds problems of interpretation or counting, it provides new counts and shows how the original software needs to be corrected to create correct counts.

Sixth, for a jurisdiction which has not been auditing all contests, both approaches require a change in **work rhythms**. Doing RLAs for all contests would require 800 - 2,600 staff hours, pulling and checking all those samples, after results are known, and fast enough so that a potential 100% hand count would be finished before the deadline for certification. On the other hand, checking ballot samples while scanning would be a change in what staff are used to, during the long periods before election day (in states which allow early scanning) and on election night or the next day. Fortunately the new procedure would start slowly as mail ballots begin arriving, for staff to get accustomed to it.

Seventh, both approaches use **software**. RLAs even use online software to do statistical tests at each step, and to compare human and computer interpretations of ballots. Software is not a necessary part of an RLA, but its common use suggests that RLAs might be hard for typical election staff to do without it.

The Options to check images or use independent rescans avoid the need to go unobserved into storage rooms, and avoid complex sampling formulas by sampling a flat number or percentage of ballot sheets, so samples can be understood, explained and monitored by staff and the public.

Colorado RLAs enter human interpretations of RLA sample ballots into "colorado-rla" software,[25] which says whether they match machine interpretations. RLAs in CA, GA, PA and VA use Arlo software[26] for the same purpose. RI uses Minerva. Colorado-rla and

---

Rules when ballot counts don't match or seals broken:

MI. state-wide: Kurth and Oosting. "Records: Too many votes in 37% of Detroit's precincts." *Detroit News.* When ballot boxes don't have the number of ballots expected, from paper records, the original counts are accepted without checking.

WI. When Wisconsin finds too few ballots in the ballot box, they recount the remaining short stack of ballots and trust the result. When they find extra ballots stuffed in the ballot box, they take out a random sample of all ballots from the box, to make the numbers appear to match, without worrying about identifying the stuffed ballots. Wisconsin Elections Commission, *Election Day Manual for Wisconsin Election Officials*, July 2016, p. 101, http://elections.wi.gov/sites/default/files/publication/65/election_day_manual_july_2016_pdf_12281.pdf. And Ansolabehere et al. 2018 Learning from Recounts https://www.liebertpub.com/doi/10.1089/elj.2017.0440

[24] Maryland State Board of Elections. (2016-10-04). "Post-Election Tabulation Audit Pilot Program Report." Page 27.

[25] FAQ4 "How does Colorado perform the audit?" https://www.sos.state.co.us/pubs/elections/RLA/faqs.html links to https://github.com/cdos-rla/colorado-rla

[26] http://www.votewell.net/auditcosts.html

Arlo are open-source, but there is no way to know that an accurate version is running at any one time. There is no information about how well the online system is protected against attackers. Staff do not tell the public the human interpretations, to allow independent checking by people who understand the statistics involved.

For ranked-choice voting (RCV) or instant runoff voting (IRV) as used in many states, RAIRE and SHANGRLA software[27] can be used to check an RLA sample, with the same issues of dependence on software and lack of public observation and understanding. Trusting any sample size, to check that the right person is eliminated and the right votes were reassigned in each round of RCV or IRV, is beyond the statistical understanding of most people. A 100% analysis of verified ballot images avoids requiring trust in samples, since every round is checked by using 100% of the images.

Eighth, most states restrict access to paper ballots after the election. Therefore a post-election RLA needs **legal authority**, which not all states have. Checking ballots while scanning might be permitted with less formality, as quality control, depending on state law, though specific authority is valuable.

Checking a 1% sample of paper ballots, provides enough sample to satisfy an RLA's statistical tests on some contests, often a majority of contests. Since 1% gives a good RLA sample for at least some contests, it complies with the recommendation of the Colorado League of Women Voters to audit all contests, and include some RLAs.[28]

RECOUNTS OF BALLOT IMAGES, AT CANDIDATE EXPENSE, ARE AN OPTION

The previous Option for election staff to check images has one main advantage and one disadvantage compared to recounts of images.

First, staff can examine paper ballots to check images early, when each batch of early ballots is scanned. Thus the paper ballots are examined when they are reliable, before they are subject to the alterations and losses while ballots are in storage, which have happened regularly throughout our history. Recounts happen after final results appear. If recounts look at paper ballots, these have been stored until the recount. If at all possible, paper ballots do need to be checked as they are scanned, even if later analysis will be done by candidates.

Second, if independent analysis of ballot images is done at the request and cost of candidates or parties, the government saves that cost while increasing public trust. Candidates can analyze as thoroughly and fast as they wish. Candidates and their supporters will have greater trust in recounts they do, than recounts done by officials who said they lost. Several states don't let candidates request recounts at all, and in the others, hand-counting or rescanning are the two approaches usually available. Those approaches are costly, especially hand-counts, and are done by election staff. NCSL has a summary of recount laws, and CEIMN has more data on state recount rules. FairVote has analyzed the 31 state-wide recounts in 2000-2019.

For 100% recounts, people use every ballot from storage. If electronic files were checked during scanning, a recount can check each paper ballot against its electronic record, or check a sample (which adds sampling risk), or count the verified electronic files instead.

Independent recounts of ballot images by outside groups cannot automatically change outcomes. Sponsors need to bring any evidence they find to a judge. They will have to show which ballot images they interpreted differently from the election software, and address doubts and counter-arguments from opposing counsel.

In current recounts done by election officials, there have been concerns that candidate representatives can be aggressive and intimidating to recount workers.[29] That approach is harmful, even though it is understandable that candidates paying for a costly recount want to gain every ambiguous vote. Recounts could go better if there is an Option for candidates' staff or contractors to recount the electronic images of ballots, instead of the paper ballots, while also checking image accuracy. The candidates can privately find which votes really may be ambiguous. If there are too few, they can give up. If there are enough ambiguous votes, a candidate can present them to election officials and/or a court, without arguing over every ballot.

---

[27] "You can do RLAs for IRV" 2021. https://arxiv.org/pdf/2004.00235

[28] LWV of CO, Election Security Position. 2022. https://www.lwvcolorado.org/docs.ashx?id=940583

[29] Carter Center 2020. **Georgia Risk-Limiting Audit/Hand Tally** p.31 "political party monitors proved to be a source of tension, hostility and even intimidation. added to the stress of the hand tally for all involved and likely affected the speed and accuracy of both counting and data entry... In four cases, disruptions were sufficient to stop auditing temporarily."
cartercenter.org/resources/pdfs/news/peace_publications/democracy/georgia-audit-final-report-033121.pdf

Candidates are likely to present only the votes which can be re-interpreted to benefit them. Winning candidates whose win is challenged can take the same approach to find other votes to benefit themselves. This costs money, but probably less than sending representatives to advocate for them in traditional recounts.

Draft wording to authorize ballot image recounts can be added to an existing state recount law with minor changes. The suggested text provides ballot images under a non-disclosure agreement without redaction, since redaction takes weeks, missing the deadlines to correct results.[30] The draft also lets candidates examine paper ballots, which stay under the control of election officials. It lets candidates choose an adequate sample of paper ballots to check accuracy of ballot images. Candidates could skip the examination of paper ballots if offices have already checked a sample of paper ballots while they were scanned, before they went into storage.

INDEPENDENT RESCANS ARE AN OPTION

Instead of checking a sample of ballots, another approach is to get an independent scan of all ballots, using different machines and software for an independent tally of these new scans.

The rescans can be done on a flow basis, before each batch goes into storage, so they have the same advantage as the Option to check ballots before they can be altered in storage. The rescans should not have the same errors as the initial election scans, so they will catch hacked images, unless a well-resourced hacker or insider can hack both systems. Faint marks can be caught if the rescan is higher resolution and/or more sensitive.

If the rescans are trusted, no one needs to look at sample ballots until after the polls close, and totals don't even need to be created until then, reducing the perception of leaking results early.

The rescan needs to release hash values of the image files for immediate distribution and publication. How do candidates and the public know the scanner doesn't fabricate image files with a predetermined outcome and release that hash value? The issue is whether hackers can hack both the official election system and the rescan system. If they can, then we need to go back to the proposal of checking some paper, randomly chosen after the hash value is published on each batch, so later changes can be detected.

Fourteen counties of Florida, with a third of the state's population,[31] pay Clear Ballot to do this rescan and reanalysis, after the original scans and tallies were done by ES&S or Dominion. Under Florida law the scanned images are public, though getting copies can be expensive. It is probably one of the most secure approaches in the country, since they re-check every contest on every ballot (and find a handful of slightly different tallies). However someone could hack both systems, which are in the same offices.

Some counties in New York have Clear Ballot rescan stored ballots some time after the election. The Vermont Secretary of State has them do the same for 6 towns. The volunteer Elections Transparency Project does so for Humboldt County, California and releases the ballot images for anyone else to analyze. Connecticut's Audit Station rescans sample election districts (precincts) in some towns and displays to staff ambiguous images & interpretations (and other ballots as desired).[32] All these provide independent checks on many parts of the process, though they won't catch changes in ballots during storage.

The reanalyses which Clear Ballot has done in CO, FL, MD are not fully independent, because Clear Ballot depends on barcodes created by the election system on some ballots. When voters use machines to mark ballots (ballot marking devices, BMDs), the machines encode votes in a barcode on each ballot, as well as printing text.

BATCH AUDITS ARE AN OPTION

Most states audit some or all contests by retallying a percentage of batches. A batch may have votes from a precinct, voting machine, drop box or a group of mailed ballots scanned at the same time. The number of batches in batch samples is usually too small to audit all contests, though the batch samples may give confidence for large contests with many batches and wide margins. Reliability for each contest depends on the number of sampled batches which contain that contest, not the number of ballots in them. Some local

---

[30] San Francisco redacts personal names which some voters write on the ballots, to certify marks where they changed their minds while marking the ballot. They release images 4 weeks after election day, (in Final report 12/1/2020),

[31] Florida counties of Baker, Bay, Broward, Columbia, Hillsborough, Indian River, Leon, Levy, Nassau, Palm Beach, Pinellas, Putnam, St. Lucie, Suwannee, according to "Levy County announces audit plans". *Levy County FL Elections*. 2021-05-03

[32] Antonyan et al., 2013, Computer Assisted Post Election Audits

contests, like school or city council districts, are missed entirely by batch samples, or these local contests appear in fewer batches than the overall sample.[33] The table below shows confidence for different numbers of batches in a sample and winning margins.

The purpose of hand-tallying a sample of batches is to find if election machines overcounted a winning candidate or under-counted others, creating a false winning margin. When a hand tally for the winner is lower than the machine tally, or hand tally for others is higher than the machine tally, auditors need to check more batches, to see if the winner's lead evaporates after tallying more batches.

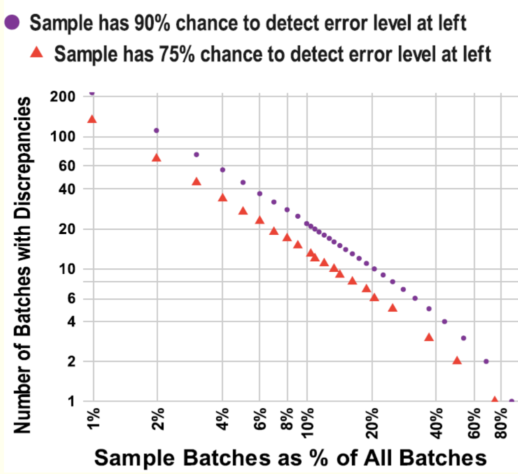
Simple analysis of batches

In very close contests, discrepancies in one batch are enough to flip the outcome from one candidate to another. For example 50 flipped ballots in a batch can cancel a lead of 100 votes, without suspicion. No one tracks whether a batch came from a drop box in a neighborhood strongly supporting or opposing that candidate, so no one knows what to expect in each batch.

Figure 4 is calculated in the "N of batches" tab in a spreadsheet. It shows the sample size needed to detect different levels of "bad" batches where the audit count, after careful review disagrees with the election machines' totals for that batch.

- If there is one batch with an inaccurate original count, a random sample needs to contain 90% of all batches to have a 90% chance of detecting the single miscounted batch.
- If two batches were originally counted wrong, a sample needs to contain 68% of batches to have 90% chance of detecting one of the two miscounted batches.
- If 44 batches were counted wrong, a 5% sample has 90% chance of detecting at least one of those miscounted batches.
- If 209 batches were originally miscounted, a 1% sample is good enough to have 90% chance of detecting at least one miscounted batch.

Figure 4. It Takes Large Sampling Rates to Detect Rare Discrepancies



An average congressional district had 350,000 votes in the 2020 general election. If jurisdictions keep ballots in batches of 200, the congressional district has 1,750 batches.

A 1% sample has 90% chance of detecting problems if there were 209 miscounted batches (these are 12% of 1,750 batches). It would take more miscounted batches than that to undermine the outcome in contests with wide margins of victory, such as many congressional and state-wide races. So the 1% sample gives confidence in these outcomes. A 2% - 5% sample would give even more confidence.

The following sections give a more detailed relationship between the number of miscounted batches and the number of individual miscounted ballots.

Threshold of small discrepancies to ignore

When auditors hand-tally a batch of ballots, most states let auditors ignore discrepancies of half a percent or so. Differences from election machines' tallies happen when either the machine or hand tally is inaccurate. Research on hand tallies has found 0.2% to 2.5% average errors in hand tallies.[34] The miscounts which are small enough to ignore under state law can be present in every batch

---

[33] California adds a batch for each contest not in the basic 1% sample, which is too few for reliability, even if California didn't omit late-arriving ballots entirely.

[34] Hand tally inaccuracies have been 0.2% to 2.5% in published studies:

- 0.8% to 2.5% in NH, in Ansolabehere and Reeves, 2004, "Using Recounts to Measure the Accuracy of Vote Tabulations: Evidence from New Hampshire Elections 1946-2002", *CALTECH/MIT Voting Technology Project*. (p.5)
- Up to 2% in CT, Antonyan et al., 2013, Computer Assisted Post Election Audits (p. 5)
- 0.5%-2.1% in experiments, in Goggin et al., 2012, "Post-Election Auditing: Effects of Procedure and Ballot Type on Manual Counting Accuracy, Efficiency, and Auditor Satisfaction and Confidence". *Election Law Journal: Rules, Politics, and Policy*. **11** (1): 36–51. doi:10.1089/elj.2010.0098

and hide under the radar. The audit looks for excess miscounts, above this threshold. These excess miscounts may be in few or many of the batches (e.g. a badly adjusted scanner, flawed software or hack). The sample needs to be big enough to find at least one batch with an excess miscount, if it exists, no matter how widespread or concentrated the miscounts are.

Table B1 is calculated from the "Batch-2pt" tab of the spreadsheet. It shows that ignoring discrepancies of 1% of a batch or half a percent, has little effect on reliability. Ignoring these small discrepancies avoids the time needed to get hand tallies exactly right.

Ceiling on inaccurate ballots per batch

The limit on discrepancies in a batch is to flip all votes from a drop box or precinct in a very partisan area to the opposite party (e.g. flip 60% of votes in a batch).

Others suggest that at most 20-30% of votes in a batch could be changed. New Mexico law & rule 1.10.23.9.A(1) "assume that the maximum margin shift in any precinct will not exceed thirty percent." A 2008 Rivest paper assumes "miscount within each precinct is at most 20% of the votes in that precinct," and cites McCarthy et al., 2007 which says, "Miscounts that shift vote totals by more than 20% of the total votes (i.e., 20 percentage points) in any single precinct should be sufficiently noticeable to trigger a suspicion-based challenge (non-random) audit." The Rivest and McCarthy papers were written when most batches were precincts, where people knew the history of votes there.

Suspicion-based audits are not widely available, and people don't see totals for each batch of mailed ballots, and don't know which batches came from drop boxes in partisan neighborhoods, so people don't know what's suspect.

One can imagine a bug or hack which changes any 80% Democratic batch of ballots from the mail or a drop box to 80% GOP or vice versa. This possibility shifts the most possible votes in the fewest batches. In the unlikely event that someone notices, it can look like a software overflow bug, caused by some normal update. Lest some say this is an obscure risk, much ink has been spilled on defense against obscure risks (hacked images, hacked internet return, etc.). Obscure undefended risks are targets for foreigners who want to undermine us and domestic attackers who want power, contracts or new rules.

Audit calculations can use a ceiling lower than 60%, if a state issues batch tallies labeled as to where they came from, and fast enough to allow suspicion-based retallies before the deadline for corrections. The table shows this helps about as much as doubling the sample size, and may cost less, but neither helps close contests. Batch samples are not effective for close contests. Suspicion-based audits are currently out of fashion in any case.

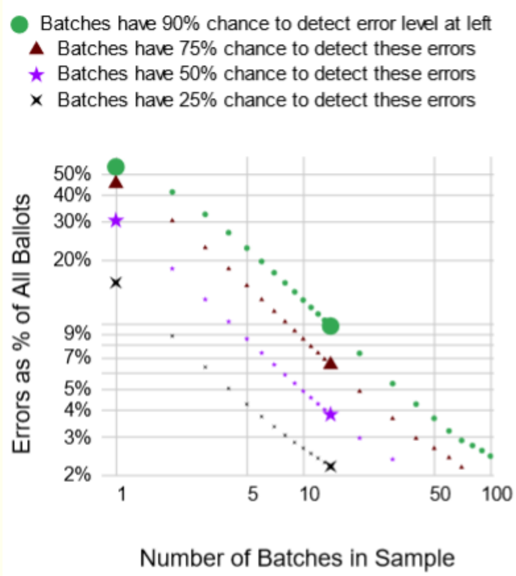
Analysis of batches using threshold and ceiling

Figure 2 is calculated in tab "Batch detail" of the spreadsheet.

Table B1 shows estimates, calculated in the "Batch-2pt" tab of the spreadsheet. For each error level shown in the left column, the right side of the table shows the minimum number of non-ignorable batches in the election which that error level could create if it is concentrated in a few batches. The right side of the table depends on the total batches in the election, size of discrepancy which auditors are instructed to ignore, and ceiling on discrepancies per batch. The left side of the table shows confidence that an audit will catch one of these batches which has excess discrepancies, above the ignorable threshold.

An average congressional district (CD) had 350,000 votes in the 2020 general election. If a jurisdiction keeps ballots in batches of 200, the CD has 1,750 batches. A 1% sample is 18 batches (spread among all jurisdictions in the CD). A 1% sample has 90% chance of detecting a miscounted batch if 208-221 of the 1,750 batches had miscounts, assuming discrepancies of 0.5% to 1% are ignored, and the maximum shift in a batch is 60%. These discrepancies can be caused by tallying 8% of ballots wrong. This analysis

Figure 2. It takes many batches to detect small errors



- 0.2% to 0.3% in WI, in Ansolabehere et al., 2018, "Learning from Recounts". *Election Law Journal: Rules, Politics, and Policy*. **17** (2): 100–116. (pp.111, 113) doi:10.1089/elj.2017.0440
- 0.2% to 0.8% in CO, in Colorado Sec. of State, 2020, pp.7,12 "Colorado Risk Limiting Audits: Three Years In", as fraction of ballots hand-counted
- 15% in AZ recount, 2021, https://archive.ph/zvoeb

is consistent with the simple analysis given earlier. It adds the story of how many miscounted ballots create this number of miscounted batches.

At worst, wrong tallies switch votes from the loser to the winner, so tallying 8% of ballots wrong can raise the margin of victory by 16 percentage points. The 1% sample, which detects this level of problem, is enough in contests with margins of victory over 16%, such as many congressional and state-wide races.

Table B1. Batch Samples' Confidence															
Total batches containing a contest		**10**	**100**	**100**	**1,750**	**1,750**	**1,750**	**1,750**	**10**	**100**	**100**	**1,750**	**1,750**	**1,750**	**1,750**
Sample as % of batches		**1%**	**1%**	**1%**	**1%**	**2%**	**2%**	**2%**	[link to this spreadsheet](#)						
Sample: Number of batches		**1**	**1**	**1**	**18**	**35**	**35**	**35**							
Threshold: Discrepancies per batch which auditors ignore		**0.5%**	**0.5%**	**0.5%**	**0.5%**	**1.0%**	**0.5%**	**0.5%**	**0.5%**	**0.5%**	**0.5%**	**0.5%**	**1.0%**	**0.5%**	**0.5%**
Max discrepancies/batch		**60%**	**60%**	**30%**	**60%**	**60%**	**60%**	**30%**	**60%**	**60%**	**30%**	**60%**	**60%**	**60%**	**30%**
Rate of miscountedballots	Questionable margin of victory if each miscount flips vote	**Confidence in catching a batch with more miscounts than the threshold, so auditors don't ignore them** *Cannot measure rare issues if ignore small miscounts							**Number of batches with miscounts over threshold is at least:**						
0.25%	0.5%	*	*	*	*	*	*	*							
0.50%	1.0%	*	*	*	*	*	*	*							
1.0%	2.0%	10%	1%	2%	14%	*	26%	46%	1	1	2	15		15	30
2.0%	4.0%	10%	3%	6%	38%	46%	60%	84%	1	3	6	45	30	45	89
3.0%	6.0%	10%	5%	9%	54%	71%	78%	96%	1	5	9	74	60	74	149
4.0%	8.0%	10%	6%	12%	67%	84%	88%	99%	1	6	12	103	89	103	208
5.0%	10.0%	10%	8%	16%	76%	92%	94%	100%	1	8	16	133	119	133	267
6.0%	12.0%	10%	10%	19%	83%	96%	97%	100%	1	10	19	162	149	162	327
7.0%	14.0%	20%	11%	23%	88%	98%	98%	100%	2	11	23	192	178	192	386
8.0%	16.0%	20%	13%	26%	91%	99%	99%	100%	2	13	26	221	208	221	445
9.0%	18.0%	20%	15%	29%	94%	99%	100%	100%	2	15	29	250	238	250	505
10.0%	20.0%	20%	16%	33%	96%	100%	100%	100%	2	16	33	280	267	280	564
12.0%	24.0%	20%	20%	39%	98%	100%	100%	100%	2	20	39	339	327	339	683
14.0%	28.0%	30%	23%	46%	99%	100%	100%	100%	3	23	46	398	386	398	801
16.0%	32.0%	30%	27%	53%	100%	100%	100%	100%	3	27	53	456	445	456	920
18.0%	36.0%	30%	30%	60%	100%	100%	100%	100%	3	30	60	515	505	515	1,039
20.0%	40.0%	40%	33%	67%	100%	100%	100%	100%	4	33	67	574	564	574	1,157
22.0%	44.0%	40%	37%	73%	100%	100%	100%	100%	4	37	73	633	623	633	1,276
24.0%	48.0%	40%	40%	80%	100%	100%	100%	100%	4	40	80	692	683	692	1,395
26.0%	52.0%	50%	43%	87%	100%	100%	100%	100%	5	43	87	750	742	750	1,513
28.0%	56.0%	50%	47%	94%	100%	100%	100%	100%	5	47	94	809	801	809	1,632
30.0%	60.0%	50%	50%	100%	100%	100%	100%	100%	5	50	100	868	861	868	1,750
Source: [spreadsheet](#) . Note: When there are only 10 batches, a whole number of them have miscounts, so chance of catching them is 10%, 20% etc.															

Here's an example of how this table calculates the number of batches with miscounts, which a sample tries to detect. You can substitute other numbers in the example.

Ballots from a small election are in 10 batches (see columns headed "10"). Auditors are willing to ignore a miscount of 1% in each batch. Suppose another 5% of all ballots were miscounted (for total miscounts of 6%; see "6.0%" in far left column). These extras (0.05 x all ballots in election) could be divided among many batches and easy to find. They are harder to find if they are divided among just a few batches, shown in the table. There's room in each batch for another 59% of the batch to have been miscounted: (1% + 59% = 60% ceiling in each batch). So there's room for this many non-ignored miscounts per batch: (0.59 x number_of_ballots/batch). The number of miscounted batches, if miscounts are concentrated and as hard to detect as possible, is:

- = (0.05 x all ballots in election) / (0.59 x ballots/batch)
- = (0.05 x ballots/batch x 10 batches/election) / (0.59 x ballots/batch)
- = (0.05 x 10 batches/election) / (0.59)
- = 0.3 batches, so it is possible as few as part of one batch is miscounted and needs to be detected. (Note the number of ballots/batch cancels out in the numerator and denominator, so the final number does not depend on number of ballots per batch or ballots in the election, just on the number of batches in the election.)

More generally, where "p" is the rate of ballots interpreted wrong by the election computers, the number of miscounted batches, if miscounts are concentrated and as hard to detect as possible, is:

- = ((p-.01) x all ballots in election) / (0.59 x ballots/batch)
- = ((p-.01) x ballots/batch x B batches/election) / (0.59 x ballots/batch)
- = ((p-.01) x B batches/election) / (0.59)
- = ((p-.01) x B / 0.59), rounded up to the next whole number. This number of batches are miscounted and need to be detected. (Note the number of ballots/batch cancels out in the numerator and denominator, so the final number does not depend on number of ballots per batch or ballots in the election, just on the number of batches in the election.)

An election with 20,000 ballots in batches of 200 has 100 batches. So does an election with 40,000 ballots in batches of 400. In both cases an error rate of 6% would mean that at least 0.05 x 100 = 5 batches were miscounted by more than a percent, and need to be detected by the batch sample. The result *does* depend on whether you ignore 1% or some other number, and whether 60% or some other number might be switched without being otherwise noticed. You can change those numbers in the example or the spreadsheet and find new results.

Chance that batch sample can detect wrong outcome

For typical US audits with samples containing 1% to 5% of all batches, the average chance of detecting a wrong outcome would be 26% to 33% in the following jurisdictions (Table B2). The lowest chances are in elections like the Allegheny County primary (Pittsburgh area), Hennepin County (Minneapolis) and Cattaraugus County (near Buffalo) where there were many small contests with few ballots, so the chance of a batch sample detecting problems was small. There is not much difference between samples of 1% to 5%, because in many contests the chance of detecting an incorrect outcome is low for any of these small sample sizes. A more expensive 20% sample would do a little better.

Table B2. Average chance (across all contests) of detecting a bad batch, when problems are common enough to flip the contest outcome

County & Election	Number of Contested Contests	Number of Ballots	Sample as Percent of All Batches		
			1%	5%	20%
AZ_Maric_2020g	201	2,089,563	41%	48%	50%
CA_Orang_2020g	166	1,546,570	13%	17%	20%
CA_Orang_2020p	48	818,021	17%	26%	28%
CA_Ventu_2020g	87	1,289,766	19%	26%	29%
CO_Adams_2020g	26	239,425	62%	81%	88%
CO_Arapa_2020g	25	373,295	37%	65%	78%
CO_Bould_2020g	33	218,692	65%	81%	87%
CO_Denve_2020g	35	1,201,712	80%	92%	96%
CO_Dougl_2020g	16	234,272	21%	35%	41%
CO_El Pa_2020g	37	382,583	44%	61%	71%
CO_Jeffe_2020g	29	398,346	55%	69%	72%
CO_Larim_2020g	24	240,986	44%	65%	71%
CO_Weld__2020g	34	196,142	18%	31%	38%
MN_Henne_2021g	49	180,000	8%	23%	40%
NY_Catta_2021g	23	37,125	4%	14%	22%
PA_Alleg_2019g	141	269,828	14%	15%	17%
PA_Alleg_2019p	139	157,264	2%	3%	4%
PA_Alleg_2020g	24	724,800	31%	41%	47%
Grand Total	**1137**	**898,790**	**26%**	**33%**	**37%**

Source: "pivot of tallies" tab in http://votewell.net/tallies.xlsx
Maximum discrepancies per batch are 60%. Auditors ignore discrepancies up to 0.5% in batch counts.

HOTLINES

Few election steps have public or party observers. To enable as much observation as possible, the times each day for processing ballots should be as limited as possible, and should be announced on the web, and followed. All steps should be webcast with high definition cameras, and staff should not know how many people are watching or who.

Observers are not usually present, so accuracy and security depend on the care and honesty of bipartisan or multi-partisan election workers. They can fear for their jobs if they report problems. The federal government and each state need effective hotlines which include both phones and anonymous web reporting tools to hear about problems and investigate them. External reporting systems are more trusted than internal ones. Hotline sponsors should summarize monthly what facts they find and what facts they lacked, especially when a lack of facts forces the hotline sponsor to close a case without a finding. Good summaries encourage workers to report more details.

Existing hotlines are limited. The hotlines at 866ourvote.org, eoldn.org and political parties do not promise confidentiality, do not report their findings and have no authority to investigate, though they give useful advice and sometimes representation. SeeSay2022.com maps reports, but has no authority to investigate. FBI offices and each US Attorney's District Election Officer only investigate after elections are certified.[35]

Businesses which make an effort to hear employee reports of problems, get a median of 1.7 reports per 100 employees per year. A quarter of these reports about businesses address accounting, auditing, financial or business integrity, which would be particularly important if election workers report problems at those levels. Half of the reports about businesses were found to be substantiated upon investigation, and some of the other half were partially substantiated.[36] "Research has demonstrated a strong correlation between report volumes and positive business outcomes," because good reporting systems solve problems before they get out of hand.[37]

If election workers report at the median level, most jurisdictions, even with hundreds of temporary workers, could not support a local reporting system for problems and good investigations, so the state and federal levels are more appropriate, and may be more trusted by workers and the public.

The fraction of voters "very confident" that "results of elections are counted fully and reported accurately" was abysmal at 27%[38] even in 2019, a year before the 2020 election controversies. The goal needs to be measuring problems, solving them and announcing the solutions.

QUALITY CONTROL FOR SCANNERS

A benefit of quality control while scanning is to correct errors before results are released. Officials, candidates and the public are happier when they have correct results in the first place. Correct results also avoid 100% hand counts which RLAs require when they find errors later. Maryland decided not to use risk-limiting audits (RLAs), because of the high sample size on close contests and chance of 100% hand tallies.[39] When quality control finds problems, it calls for rescanning on a better scanner or a repaired scanner, and re-checking, which are easier than a complete hand count.

The main unresolved risks in all Options for tabulation audits above are accuracy of ballot scanning, and the storage of ballots and electronic records. This section of the paper shows how quality control samples for election scanners address these risks.

---

[35] US Justice. (2017). *Federal Prosecution of Election Offenses*. US Justice Department. http://www.votewell.net/a/doj-elections.htm

[36] Navex. (2022). *2022 Risk & Compliance Hotline & Incident Management Benchmark Report*. (pp. 8, 19, 29) https://www.navex.com/en-us/resources/benchmarking-reports/2022-risk-compliance-hotline-incident-management-benchmark-report

[37] Navex. (2021). *2021 Risk & Compliance Incident Management Benchmark Report*. (p. 10) https://www.navexglobal.com/en-us/resources/benchmarking-reports/2021-risk-compliance-incident-management-benchmark-report?RCAssetNumber=8644

Stubben, S., & Welch, K. (2020). Evidence on the Use and Efficacy of Internal Whistleblowing Systems. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3273589

[38] Knight Foundation. (2020). *The Untold Story of American Non-Voters*. (p. 11). https://the100million.org/

[39] Maryland State Board of Elections. (2016, October). *Post-‑Election Tabulation Audit Pilot Program Report*. https://www.elections.maryland.gov/press_room/documents/Post%20Election%20Tabulation%20Audit%20Pilot%20Program%20Report.pdf

It is *not* always possible to see directly when skilled crooks bypass locks, seals, security cameras and chain of custody records to change stored paper ballots.

We *can* detect changes in stored ballots by (a) using a quality control sample to ensure we have accurate electronic copies of ballots before the ballots go into storage, (b) preserving the electronic copies with digital signatures, and (c) using the verified electronic files to check if paper ballots were changed during storage. The verified electronic files give us more accurate initial results, and let us recover true copies of ballots if needed.

Ballots can be, and usually are, safely under bipartisan observation, which prevents changes, until the ballots are scanned. Places which store mailed ballots unscanned, out of sight of observers and staff, lose this benefit of continuous bipartisan observation. They will have no way to check whether ballots are changed in storage.

Digital signatures and hash values

Election machines scan ballots to create electronic images and cast vote records (CVRs). The machines run an algorithm on each file to calculate a digital signature or hash value of each file.

Dominion says, "Once the polls have been closed, the ImageCast tabulators encrypt all of the results files prior to transmitting them back to EMS. SHA-256 *hashes* are used for all data integrity and verification. Should an intrusive process or altering of any file occur, hash values will be, in turn, altered as well. Any presence of an intrusive process will be detected, as the hashes of any altered data will not match the value initially determined."[40] Dominion's hashes are not limited to ImageCast tabulators. They say, "All electronic records are *digitally signed*, all election records and database files are encrypted."[41]

ES&S similarly creates *digital signatures* (which contain hash values) when it creates images, "ES&S software digitally signs every cast vote record and its corresponding image files when they are created… encryption and digital signing for all data-in-transit using cryptographic modules that meet the Federal Information Processing Standard 140-2 (FIPS 140-2)."[42] ES&S also says its precinct scanners create files which are "digitally signed and encrypted at poll close."[43]

CISA says a hash value is "a fixed-length string of numbers and letters generated from a mathematical algorithm and an arbitrarily sized file such as an email, document, picture, or other type of data. This generated string is unique to the file being hashed" (US CISA, 2021). Digital signature is a related term. It is a hash value, encrypted with a numeric key from the scanner or election office. It can be decrypted later, using a paired key which the public has, to find the hash. A digital signature can be mistrusted if someone forges a hash and gains access to the election office's key to encrypt the forged hash, or uses a new key and posts the corresponding decryption key in place of the original decryption key. Posting each hash on a secure time-stamped archive, is a way to create trust, with a digital signature from the time stamp authority.[44]

Once the hash value for an electronic file is calculated and published, anyone who has a copy of the file later can run the same algorithm. If the new result matches the original value, the electronic file is unaltered. This system makes storage of electronic files more testable than storage of paper files.[45]

For batches processed before election day, it is not possible to release the computer files immediately. A hash value needs to be released to ensure computer files are not changed later. The air-gapped election equipment can have a printer to print each hash value, possibly with a bar code for easy data-entry into other computers. To avoid doubt later about what the original hash value is, copies should go immediately to all observers, reporters, safes under control of different staff, and a time-stamped online archive.

---

[40] P.56 of Dominion and Georgia. Master Solution Purchase and Services Agreement 2019. https://gaverifiedvoting.org/pdf/20190729-GA-Dominion-Contract.pdf

[41] Dominion Voting Philadelphia Bid#1903107, 2019. P.4 of files7.philadelphiavotes.com/announcements/Dominion_-_Redacted.pdf

[42] ES&S Responses to State of New Hampshire Questions. 2019. Pp.27-28 of https://sos.nh.gov/media/2wydqgkb/es-and-s-response-to-nh-questionnaire.pdf

[43] "DS200® Precinct Scanner and Tabulator" https://www.essvote.com/storage/2020/12/DS200-Security-Bulletin.pdf

[44] Waldvogel, M. (2021). *PGP Digital Timestamper timestamped archive*. https://github.com/zeitgitter/pgp-digital-timestamper-timestamped-archive

freeTSA.org. (2022). *Time Stamp Authority*. https://www.freetsa.org/

https://en.wikipedia.org/wiki/Trusted_timestamping

[45] San Francisco's answers on their hash codes http://www.sfelections.org/results/20160607/data/SHA512_FAQ.pdf

Election machines do calculate hash values at the time they create ballot image files, and should release those values immediately so staff and observers can check at any later time whether the file is unaltered. A digital signature for each ballot cast in person is not wise, since its creation time reveals the order in which ballots were voted. A hash for each ballot and a hash for the list of hashes or a hash for the whole file, timestamped after the last ballot is deposited, is safer.

During election processing, auditing and recounts, it is important to check that the hash values of the files being studied match hash values published when the images or CVRs were first created. This means having the files in the same format used to create the original hash values.

Election machines scan ballots. From the electronic scanned files, election machines calculate, but do not reveal, the hash value or digital signature. Election staff need to get the hash value, print it on paper for each staff member & observer present, and post it online in a time-stamped archive. Computer experts will trust the archive, and the public will trust paper, so both are useful.

Role of software for quality control

Quality control software checks all scans to see if they have stretched images, unreadable barcodes, spurious black or white lines from folds, bad sensors, etc.[46] or other issues which software can detect. Good software works around minor issues, and points out the rest for correction.

Quality control software can also detect images which repeat within a batch or across batches. Scanning ballots twice can happen accidentally or to fix problems, and is hard to notice if other ballots were missed, if the total number of ballots is close. Hand marked ballots fill in the targets with slightly different pen strokes, which identify duplicates, even when two voters choose the same set of candidates. Machine-printed ballots are much harder to tell apart than hand-marked ballots, though a batch of ballots which repeats later in the same order, with the same write-ins, is worth investigating. RLAs and batch percentage audits look at so few ballots, they will not notice duplicate ballots.

Manufacturing and service industries test their systems both before production starts, and also during production. Election staff use logic and accuracy tests in the weeks before elections to check hardware and software. They do not always do quality control checks during elections, and need to.

Role of humans for quality control

Besides 100% quality checks by software, quality control also includes humans comparing a sample of paper ballots to their electronic records, to detect issues which software misses, before the paper ballots go into storage.

Humans compare each random ballot to its freshly scanned electronic record. Ballots scanned in polling places may need rescanning centrally so that ballots and electronic records will be in the same order, which makes this one-to-one matching easy. Options are in Table E2 below.

If any electronic record does not match its paper ballot, staff need to find and correct whatever caused the flaw (scanner, software, feeder, etc.). Then the office rescans any batches affected by defects (all batches if necessary), re-samples them, re-tests them.

Each defect at this quality control stage needs to be resolved, until we see no differences between the electronic files and the paper ballots. This stage ensures initial election results are as accurate as possible. These accurate files provide a baseline to detect any changes during storage. Copies of the verified electronic files go into separate safes, of multiple elected officials or judges, to reduce the chance an accident or hack can harm all copies at once. The files must stay in the same format used to create the original hash values, so hash values will match. Safes would not protect the files if all officials cooperate in changing them, which might happen in machine politics or a company town. In that case the published digital signature will show that no copy matches the original, and an outside judge will have to decide what to do.

Human errors. Colorado has measured inaccuracies in human processing. 90 (1.1%) of the 8,306 ballots sampled in Colorado in the 2020 general election had inaccurate classifications by the manual audit, in the view of the Secretary of State's staff. These included 0.3% of the sample where programming was inaccurate in mapping contest names, 0.2% where SOS staff thought auditors counted to the wrong ballot (by mistake, or because the ballot file had more or fewer ballots than the CVR file), and 0.6% where SOS staff

[46] Walker, N. (2017). 2016 Post-Election Audits in Maryland. US Election Assistance Commission. https://www.eac.gov/sites/default/files/event_document/files/day1-2017-02-13-09_15-03-natasha-Post-Election-Audits-final.pdf

Gideon, J. (2005). *Hart InterCivic Optical-Scan Has a Weak Spot*. Voters Unite. http://www.votersunite.org/info/yakimaproblemreport.asp disagreed with county auditors and thought there was no discrepancy. For remaining discrepancies where SOS staff agreed with auditors, 0.2% of the sample had incorrect adjudication in initial processing, 0.1% had the voting software count a mark crossed out by a voter, and 0.04% had ambiguous voter intent.[47]

Counting ballots. Counting in a stack of ballots to reach the sampled ballots has a low error rate. Colorado has found that the wrong ballot is reached one in 300-600 times (0.2-0.3%). It is possible some of the counts were accurate, and the different destination is due to an extra ballot or CVR in their files. Weighing is another approach, if ballots have not picked up significant dirt, coffee, etc.[48]

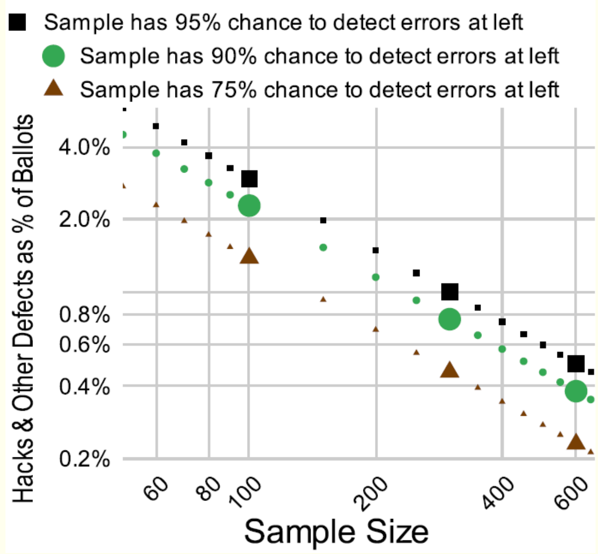
Cast vote records or images

Sample comparisons by humans can indeed compare paper ballots to either CVRs or ballot images.

Checking CVRs by humans needs more time than checking images, but both are feasible. The software which interprets ballots into CVRs has many more ways to create inaccurate CVRs than the software which stores images. Resolving these extra CVR inaccuracies takes time, and would not help identify inaccurate ballots later.

Checking done by software can be done on images, and often cannot be done on CVRs, which may not even be created when bar codes identifying style are unreadable. Software checking images can reveal problems such as stretching and spurious lines, which CVRs do not reveal. Checking CVRs without checking images puts the entire burden on the sample and humans.

Figure 5. Chance that sample of ballots will detect any hack or other defect



Sample sizes

When humans review a sample, confidence in these reviews is subject to sampling risk, shown in Figure 5. For example, comparing a sample of 100 ballots to their electronic records has 90% chance of catching an issue which affects 2.3% of records, and 75% chance of catching an issue which affects 1.4% of records. For more money, studying 600 random ballots has 90% chance of catching inaccuracies which affect 0.4% of records. These give useful reassurance for almost all contests, with as much or more precision than the error level in hand counts.

In table S, bigger sample sizes (to the right) can detect smaller errors, at each level of confidence (each row).

[47] Colorado SOS, Discrepancy Report, 2020, https://www.sos.state.co.us/pubs/elections/RLA/2020/general/DiscrepancyReport.pdf Percentages are calculated in the "CO errs" tab of http://votewell.net.tallies.xlsx

[48] **Weighing ballot sheets** on scale is endorsed by EAC (p.16) and *Audit Principles & Best Practices* (p.12 item 4e) .

An issue is which counting method is most accurate. Orange County CA used a scale. "election agency staff report that in their experience the scale is accurate to within one or two ballot sheets" which may or may not be adequate: Note 20 in https://verifiedvoting.org/wp-content/uploads/2020/08/2018-RLA-Report-Orange-County-CA.pdf

If batch includes different types of ballot sheets, such as computer-generated or mailed, they need to be weighed separately, with separate lookup tables to give count of ballots for each weight.

A page of 8.5 x 11 paper (20-pound bond or 50-pound offset) weighs 4.54 grams or a sixth of an ounce. The scale has to be accurate enough to make those distinctions in piles of 200 to 500 ballots (2-5 pounds), 1 part in 200-500. Weights are proportionately heavier for 24-pound paper and larger ballots. An electronic scale can be that accurate, and may be trusted more than hand-counting the piles. https://www.mt.com/us/en/home/library/know-how/industrial-scales/weighing_votes.html

Table S. Sample Size	50	100	200	300	600	1,000	2,000
Chance of detecting issue	**Errors Detectable, for Confidence at Left & Sample Above**						
95%	5.8%	3.0%	1.5%	1.0%	0.5%	0.30%	0.15%
90%	4.5%	2.3%	1.1%	0.8%	0.4%	0.23%	0.12%
75%	2.7%	1.4%	0.7%	0.5%	0.23%	0.14%	0.07%
50%	1.4%	0.7%	0.3%	0.23%	0.12%	0.07%	0.03%
25%	0.6%	0.3%	0.14%	0.10%	0.05%	0.03%	0.01%
Source: "srs" tab of spreadsheet							

For example, suppose initial election results show a margin of victory of 0.8% in the closest contest. If just 0.4% of records had a problem which moved votes from the loser to the winner, that would raise the margin of victory by 0.8%. So a margin of victory of 0.8% can be created by problems in as few as 0.4% of records.

Sample size determines how small an error the random sample will detect. Most elections have some close contests where a small error can flip the outcome. Staff need to pick a sample size for quality control on the scanners before results are in, when no one knows what the closest contest will be. The choice of sample size depends on budget and local patterns of close elections. Any checking is better than the current system of not checking storage and not checking most contests. A Bernoulli polling audit can determine its sample size as it goes, based on what it finds in the closest contests,[49] but it needs a very big sample, and it reveals the voting results, so it cannot be done when ballots are being scanned long before election day.

State laws do not specify quality control samples. Offices need to choose a sample size based on budget and how close local contests have been. Staff also estimate how many batches they will have. Each batch can be ballots from a precinct or mail bag or any ballots which are scanned together as a batch. If a small office expects to have 20 batches, a sample of 100 ballots means 5 ballots per batch. For a big office with 2,400 batches, a sample of 300 ballots means one ballot per 8 batches.

Ten-sided dice. Dice are available marked 0-9, 00-90, and 000-900.[50] Throwing all three gives a random number from 0 to 999. Throwing dice in a dice tray allows good randomness in a small space.

For a sample of one ballot per 8 batches, staff throw the dice for each batch, and choose the batch if the random number shown by the dice is under 125. To choose a ballot from a batch, staff throw the dice to select a new random number and count to that ballot in the batch. If the dice show a number bigger than the batch size, staff throw again. When each batch has less than 400 or 600 ballot sheets, staff can avoid high-numbered throws by using a 4-sided or 6-sided die for the hundreds die.[51]

The alternatives of choosing a series of random numbers before all batches are final, or using a scratch card to provide random numbers, are less trustworthy, since the computer knows what the next numbers will be and can be suspected of altering other ballot images, the ones which won't be selected.

Formulas for simple random samples. There are formulas connecting sample size, error level, and risk limit or confidence level in detecting those errors. The following formulas apply to simple random samples. Stratifying by batch, as proposed here is never worse and often better.[52]

e = *erroneous* electronic records as percent of *all* electronic records (CVRs or ballot images)

e = chance that any one *ballot sheet*, examined at random, has an erroneous *electronic record*

(1-e) = chance that *one* random ballot sheet will *miss* all the erroneous records, ie, it has an *accurate record*

---

[49] Bernoulli Ballot Polling: A Manifest Improvement for Risk-Limiting Audits, Kellie Ottoboni, Matthew Bernhard, J. Alex Halderman, Ronald L. Rivest, Philip B. Stark https://arxiv.org/abs/1812.06361

[50] Note that dice made from transparent or translucent material are easier to demonstrate as non-weighted, as matching the index of refraction is more difficult than weighting an opaque die. https://diceemporium.com/product-category/d10-dice/

[51] A 4-sided die needs to be elongated or have arched ends to roll well. https://role4initiative.com/products/set-of-7c-translucent?pr_prod_strat=description&pr_rec_id=bd348c566&pr_rec_pid=7511051239643&pr_ref_pid=7511051960539&pr_seq=uniform

[52] "**Stratification** will yield a sample that is at least as precise as a simple random sample of the same sample size." p.138 of Chapter 5 of Johnnie Daniel, 2012 *Sampling Essentials: Practical Guidelines for Making Sampling Choices* https://www.sagepub.com/sites/default/files/upm-binaries/40803_5.pdf

$(1-e)^n = r$ = risk that ***n*** random ballot sheets will *all miss* the erroneous records

That final formula can be rearranged to calculate error rate and sample size:

$1-r^{1/n} = e$ = error *rate* which a sample of **n** can detect, with only **r** chance of missing it (fractional exponent means the **n**th root).

e multiplied by total number of ballot sheets = *number* of erroneous ballot sheets which a sample of **n** can detect, with only **r** chance of missing it

$\log(r) / \log(1-e) = n$ = *sample size* needed, so the chance of missing error level **e** is only **r** (logarithms can be natural or any base, as long as they have the same base as each other).

Checking after storage

When ballot image audits (BIAs) do not use stored ballots, they can skip this step. For RLAs and batch percentage audits, auditors go into storage rooms to bring out a sample of stored ballots. Auditors also need to get the electronic files corresponding to these ballots and run the same computer algorithm which calculated the original hash values or digital signatures. If a hash value does not match, auditors need to get a copy of the file from another safe and be sure its hash value matches the original.

Auditors audit the sample ballots they pulled from storage. They also need to compare these ballots to the electronic files. If any stored ballots do not match the verified electronic files, auditors need to investigate further, and may need to replace changed ballots with better copies from the electronic files. Investigation includes checking the CVR, scanned ballot image, and their digital signatures. If the image matches the paper ballot, and the CVR is wrong from bad software interpretation (most common), this is a discrepancy which RLA calculations cover. If necessary, the sample will expand, potentially to a 100% hand count,

If an image is not a good match with the paper ballot, the first check is whether auditors counted to the wrong ballot, which can happen when there is a counting error or when the batch of paper ballots has more or fewer ballots than the electronic file, or one is out of order. If the counting is correct, investigation needs to address a potential crime of adding, removing or changing a paper ballot or electronic file. Investigation includes checking the paper ballot for signs of forgery, and checking security cameras and seals. Improvements may be needed for the future. If no evidence is found to prove what was changed, the presumption is that the paper ballot was changed during storage. The more such discrepancies there are, the more effect they have on outcomes, and the stronger the conclusion that the paper was changed, since the QC sample had a known low probability of missing discrepancies.

For 100% recounts, people use every ballot. They can check each paper ballot against its electronic record, or check a sample (which adds sampling risk), or count the verified electronic files instead.

Combination of sampling & storage risks

Most audits have sampling risk. For example an RLA may have 95% chance of catching inaccuracies in tabulation. Checking storage also has sampling risk, equal to the sampling risk in the original file verification (at time of scanning). For example it may have had 90% chance of catching an error level big enough to change a particular contest's outcome. When an RLA of that contest checks both tabulation and storage, the combined chance that the audit will catch inaccuracies is 95% times 90%, or 86%. When an RLA does not check storage, its combined chance is effectively 95% times 0%. One can also estimate and include the chance of flaws in the RLA software.

When a 100% recount retallies all ballots, and verifies them all against their electronic records, the tabulation has no sampling risk. The 90th percentile error in hand-tallying is 2.5%.[53] The overall chance of catching inaccuracies up to 2.5% is then 90% times the 90% chance of catching issues in the original file verification, or 81%.

A BIA uses software to count votes on all ballot images, plus human checking of all ambiguous cases when they are enough to change an outcome. A BIA tabulation, like 100% hand counts, has no sampling risk. The overall chance of catching inaccuracies can be more or less accurate than hand counts, depending whether error rates are higher among (a) humans counting all ballots, plus software for totals, or (b) humans checking ambiguous cases, plus BIA software.

---

[53] Goggin, S. N. (2022, October 2). *Personal communication.* "The 90th percentile error across all groups and races was 2.5%" for experiments in Goggin et al. 2012, "Post-Election Auditing: Effects of Procedure and Ballot Type on Manual Counting Accuracy, Efficiency, and Auditor Satisfaction and Confidence".

In any audits, public files let others check all steps. Most RLAs and batch percentage audits do not release detailed public data and are uncheckable. Ballot image audits are done with public data in Humboldt County, CA, FL, GA, MD and VT, and are checkable.

Comparison rules to check if election records match paper ballots

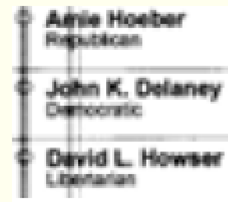
Types of mismatches between ballots and election records can happen by accident or by malicious action:

A. Faulty sensor in scanner puts white or black vertical line down all or part of one or more images. Good software checks that lines do not hide votes.

B. Stepper motors fail to advance the paper evenly under the sensors, stretching, compressing or twisting one or more images enough to make vote selections or style numbers, etc., ambiguous. Good software adjusts for many variations and reports the rest.

C. An area of one or more images is too light or dark to see vote selections or style numbers. Good software adjusts for a wide range of darkness.

D. Scanner software shifts filled and blank ovals between candidates (unlikely to be accidental).[54]



Flawed scanner put black lines on image, making overvotes

Both humans and software can see A-C. They are likely to be present sporadically in any large file of electronic records. Software needs to look for these on all ballots, not just a sample, immediately, before ballots go into storage, and before initial results are issued.

Human checking needs to look for D. While looking for D, they also provide a human check on A-C. The rule could be either:

1. If a black or white line, stretch, twist, or variation in darkness is visible to the human staff, they call it a flaw and get scanner cleaned or replaced. This can save time by correcting problems before they get severe.

2. Or: If a black or white line, stretch, twist, or variation in darkness exceeds reference samples provided by the jurisdiction, they call it a flaw and get scanner cleaned or replaced.

Checkers can avoid getting in a rut, since they only check one or a few ballots in each batch. Then they do several minutes of other activities: putting away the previous batch, opening, scanning & sampling the next batch.

Screen projection. Comparing images to ballots can be done by projecting both on a screen. An opaque projector can be used for the paper ballot, so it is not filtered by a computer projector. Checking images visually, while they are projected on a screen, depends on human concentration. It's not a rapid, rote repetition, since there are breaks between examinations of images. Breaks occur for counting ballots, scanning, handling boxes, etc. Like any human effort, it could still miss things.

Data entry. Another approach is for humans to enter votes from the paper ballot into a computer, have the computer compare the humans' entries to the voting system's interpretations, then check if any discrepancies are due to image errors. Rhode Island found that verifying 10 contests from a ballot sheet in this way took a minute (62 seconds[55] for a team of 5, so 5.1 staff minutes, though 4 would be enough). This is the approach used in ballot-comparison RLAs in California and Colorado. It is at risk from software errors, which could hide mismatches, accidentally or on purpose.[56] This is more than the time which would be allowed for visual comparison, 30 seconds for each side of a ballot sheet for a team of 2, which is 2 staff minutes.

Ballot & image should be projected on a screen even with data entry, so in-person observers, remote observers on webcams, and any supervisors nearby can compare for themselves, since the comparison software is itself hackable. Interpretations would not be projected, to avoid biased results. As discussed below, use of webcams, where staff do not know who or how many are watching, supplements in-person observing which is rare in most jurisdictions.

---

[54] Ballots for Dominion Voting Systems have blank ovals printed in red ink, and Dominion scanners drop out red from images, unlike other vendors where blank ovals are clearly present unless the image is flawed.

[55] Note E8 on p.59 of *Pilot Implementation Study of Risk-Limiting Audit Methods in the State of Rhode Island*. 2019. https://verifiedvoting.org/wp-content/uploads/2020/07/RI-RLA-Report-2020.pdf

[56] An example of different software hiding mismatches is hash-checking which hides the situation where no comparison hash is provided. https://zetter.substack.com/p/votings-hash-problem-when-the-system?s=r

	Table E2. Multiple Approaches to Check if Electronic Records Match Paper Ballots							
	Staff hours per 1,000 ballot sheets							
	1.0	*1.4*	*0.8*	*1.0*	*1.8*	*44*		
	Ballots scanned centrally (eg. mailed ballots)	**Ballots scanned in polling places, 5 approaches**					**Jurisdictions without ballot images, 2 approaches**	
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
1	When each batch is scanned, take the steps below to examine a 1% sample of ballots[A]	Rescan all polling place ballots centrally. Then treat like (a)	Rescan random 10% of precinct ballot boxes (if find problem, best to shift to (b) and rescan all)	Before scanning print random ID[G] on every ballot	Plan to find any precinct ballot in CVR file (hard if few contests or many straight party voters)	Publish tallies of all ballot boxes after polls close, then pick sample of 18 boxes.	Require that images be saved wherever the scanner can do so. Then approaches on the left can be used. Otherwise take steps below to check files of Cast Vote Records (CVRs). For (g) rescan polling place ballots centrally.	
2	Central office publishes hash values[B] of image files (and/or CVRs, as available) on paper to anyone present, and online in a time-stamped archive, to deter and detect any later file alterations. Hash values and sealed electronic files are put in safes for backup.							
3	Throw 10-sided dice to select 1% sample[A]		Throw 10-sided dice to select 10% sample of each sampled batch.[F]	Throw 10-sided dice to select 1% sample of ballots[A]		Throw 10-sided dice to select sample of ballot boxes	Throw 10-sided dice to select 1% sample of ballots[A]	Throw 10-sided dice to select 1% sample of batches (at least 18 batches in small counties, though this will have trouble detecting scanner problems affecting few scanners).[A]
4	Count to those random numbers in stack of paper ballot sheets fresh out of the central scanner.[C] .			Count to those random numbers in pile of ballot sheets from ballot box		Tally those ballots by hand.	Count to those random numbers in pile of ballot sheets.	After polls close, hand-tally all contests in each batch in sample.
5	Count to the corresponding ballot images in image file.			Electron-ically search for their random codes in image file	Search for their pattern of answers in CVR file. Check if image matches paper	Compare hand tally to previously published machine tally.	Compare sample of ballots to their Cast Vote Records (CVRs)	After polls close, sum CVRs for every batch, compare to official tallies & hand tallies
6	If any part of the image differs from the paper ballot, (a) report differences, (b) fix the scanner or use another,[D] (c) rescan, and (d) re-post the hash. It is important to throw the dice again to get a new sample to test the corrected file.[E] Using a new sample avoids checking the same ballots, so no one can suspect that staff just corrected the few ballots in the sample. Staff can repeat these steps until images are accurate.					If any vote differs between the paper ballot and CVR, (a) report differences, (b) investigate to see if they can be fixed (poor scanner or software creating CVRs), (c) fix problem, (d) re-process, (e) re-post hash, (f) re-throw dice, until CVRs are accurate; or (g) hand-tally all ballots.		
7	Investigate any differences, to see if they affected preceding and later ballots, find causes, and find preventive measures							

Table Notes

[A] **Sample size** needs to be chosen in each state. It is a trade-off between cost and accuracy. In small counties (or in towns administering their own elections) the county's (town's) sample should have at least 100 sheets.

[B] A **hash value** is a fixed-length string of numbers and letters generated from a mathematical algorithm and is unique to the file being hashed, as defined in CISA ST04-018. https://www.cisa.gov/tips/st04-018

[C] Ballot sheets can be counted by hand or by precision scale.

[D] If a **discrepancy** happened because the scanned image lost faint marks, may need to change to gray-scale, color, or higher resolution.

[E] **After finding one problem,** the office may want to sample more than 1% on re-do.

[F] 10% sample of 10% of precinct boxes gives 1% sample of all precinct ballots. It avoids rescanning most precinct ballots. The clustering makes this somewhat less reliable than simple 1% sample, so may want to select 15% to 20% sample *within* each sampled box, to give overall 1.5% to 2% sample of precinct ballots. Statisticians calculate effective sample size by comparing intra-class and inter-class variability, and can optimize the right mix of sampling ballot boxes and sampling within boxes, depending on cost of each.

[G] **Random ID on ballots:** Some people have proposed printing individual code on each precinct-scanned ballot to tie the paper ballot to its scanned record. A code could be printed on each ballot inside precinct scanners capable of doing so. If printed inside the scanner it should be a rare ink color, like purple, green or brown, and printhead should be physically to left or right of any area where votes are marked, so it cannot print votes. Trusting software to control its marks would add another target to hack. Such a code needs to be random. A sequential code would let someone who knows the order of voters find a voter's sequential ballot.

Reports. Staff need to report discrepancies even if they can be corrected. Immediate reporting ensures that problems are not forgotten or hidden, and that other jurisdictions see the kinds of problems to watch out for. Each state, or the nation, needs a clearinghouse to post audit reports for comparison. Then people see if there are patterns, such as errors in sheriff races or in small towns, and address them. An important issue is to monitor if hacked ballot images start appearing in actual elections, and where they appear, so they can be caught nationally.[57] Each jurisdiction also needs to issue a summary report after processing all ballots.

Checking centrally or at precincts. Ballots scanned at precincts can in principle be checked at the precinct or centrally. Bringing them to a central office is easier to supervise and audit, at the location where all the non-precinct ballots were audited, and where centralized observers can watch them.

The option for auditing at the precinct may be useful for a remote precinct, where local observers can get to the precinct more easily than to a central location, especially if some of them mistrust the central location. Rescanning the paper ballots at the end of the day at the precinct or centrally will create electronic records in the same order as the ballots. The precinct scanner may be slower than a central scanner. At 6 seconds per ballot sheet, 0.1 minutes, 300 ballot sheets will take 30 minutes, plus 4-11 minutes for the same extra audit steps which would be needed centrally. Precincts often have 4-5 staff all day, so two could handle the scanning while the others fill out forms to close the polls. Or two fresh staff can arrive at the end of the day for the auditing.

Whether rescanned centrally or remotely, the rescan at the end of the day which puts ballots and electronic records in the same order, also gives another check. Any difference between totals accumulated during the day and the rescan needs to be resolved.

When to check paper ballots

Checking paper ballots before they can be altered. Checking individual ballots against images or CVRs, as they are processed, has advantages. Checking while scanning is the most reliable time, since ballots have stayed under control of bipartisan teams. Paper ballots have not yet gone into storage where insiders and criminals can and have been known to alter them.

All seals, locks and cameras can be bypassed. The federal government has standards for seals used in shipment, none for storage.[58] Shipment seals must resist tampering for 30 seconds.[59] Manufacturers must have third party testing and never make seals with the same design and number. Buyers must check the manufacturer's security as well as their own.[60] The seal logs themselves must be kept securely, with an independent chain of custody.

---

[57] Clearinghouses need to be factual and not draw overly general conclusions. CISA and other agencies lost credibility 11/12/2020, when they felt the need to endorse the election before audits were done, "election was the most secure in American history." https://www.cisa.gov/news/2020/11/12/joint-statement-elections-infrastructure-government-coordinating-council-election OSET said this just meant that CISA saw less "suspicious network activity" than 2008, 2012 and 2016. https://steveschneider.substack.com/p/most-secure-election-in-american

[58] **Guidance for storage:** EAC, Best Practices: Chain of Custody (July 13, 2021) https://www.eac.gov/sites/default/files/bestpractices/Chain_of_Custody_Best_Practices.pdf

Luther Weeks, "few or no protections in place to prevent a single person from accessing the ballots" https://ctelectionaudit.org/2022/ObservationReport2021Nov.pdf

Bernhard, Benaloh, Halderman, Rivest, Ryan, Stark, Teague, Vora, and Wallach. "No state has laws or regulations to ensure that the paper trail is conserved adequately, and that evidence to that effect is provided." https://mbernhard.com/papers/voting-sok17.pdf

Stark, An Introduction to Risk-Limiting Audits and Evidence-Based Elections, for the Little Hoover Commission https://www.stat.berkeley.edu/~stark/Preprints/lhc18.pdf

> *Check that the number of ballots ... [is consistent at each step]*
>
> *Check signature verification...*
>
> *Review surveillance video...*
>
> *Check each seal against its photograph...*
>
> *Record any discrepancies.*

[59] FF-S-2738A March 30, 1999, Federal Specification, Seals, Antipilferage, http://everyspec.com/FED_SPECS/F/FF-S-2738A_25291/

[60] Dept of Homeland Security, *User's Guide on Security Seals for Domestic Cargo* January 2007 p/4-3, 6-9. https://www.hsdl.org/?view&did=235928

Checking while scanning spreads out the workload, instead of leaving it to the rushed period between election day and canvass. It is also the cheapest time to do it, since the ballots are at hand, and don't need to be retrieved from storage. Checking later is possible but less reliable.

The biggest workload is to check the many ballots which arrive on election day. There are two approaches. The first approach is to bring in a second shift with fresh workers and supervisors for the evening. This has an advantage in finding any problems before preliminary results are released. The second shift also lets the regular shift the next day move on to other post-election tasks. In an atmosphere of public suspicions, publicly checking ballots before they go into storage will allay some of those suspicions. The second approach is to store ballots overnight and either accept the loss of trust in the ballots or let multi-partisan observers keep watch on the storage. Republicans in 2000 guarded a warehouse of voted ballots in Florida and had state police lock up ballots in New Mexico.[61] Parties monitor overnight ballot storage all over the world.[62]

Preventing leaked results. In most states, processing starts before election day for mailed ballots. Examining a sample of ballots at that early time reveals them to staff and observers. This gives almost no information about total votes, because so few ballots are seen, and early ballots are not representative of final results. Early processing has always let staff and observers see individual ballots while adjudicating, re-creating, or scanning them. Staff and observers need to be sworn to keep ballots secret and to face penalties if they don't.[63]

The precision of leaked results from an early sample is limited, because of sampling variability, and the fact that early voters differ from later voters. For example, by the time 500 ballot sheets have been sampled, observers know the results of the sampled *ballots* exactly. There is an error range of plus or minus four percentage points between that sample and the unknown true tally result of the sampled *batches*, and a further difference from the later voters. The early sample can easily show results up to 10 percentage points different from the final result, so it doesn't mean much. A closely divided county will not learn anything. A strongly partisan county already knows its result.

When to check batch totals. In approach (h) getting batch totals for a representative sample of early batches may reveal too much before election day, so batch tallies should probably happen after polls close, which does leave early ballots vulnerable to alteration in storage.

States which do not allow processing before election day. If mailed ballots are not allowed to be processed before election day, or election night, it would be wise to use tamper-evident envelopes which reveal attempts to open them by steam, water or chemicals. There must be excellent locks,[64] controlled by opposing parties, and perhaps observers around the clock as mentioned above in Florida, New Mexico and around the world. Designing security cameras to minimize hacking, and locks to minimize intrusion, is beyond the scope of this paper. If locks and observers are so strong that all candidates or parties trust them, then sampling while scanning will not be necessary. After polls close, staff can release batch totals as soon as each batch is scanned, with or without hash values. They can select and examine a sample of paper ballots soon, but not immediately, and they may design the sample to handle the closest contests as in an RLA.

---

[61] Cobb, Sue (2016-10-17). "The 2000 Presidential Election – The Florida Recount". *Association for Diplomatic Studies & Training* and Baker, Deborah (2004-10-31). "ABQjournal: Contentious 2000 Election Closest in N.M. History". *Albuquerque Journal*.

[62] The Organization for Security and Cooperation in Europe (OSCE) has called for independent foreign officials to sleep with ballots, and lets parties do so: International Crisis Group (ICG) 1997. "Municipal Elections in Bosnia and Herzegovina". *RefWorld*. "OHR SRT News Summary, September 7, 1998". *Office of the High Representative (Bosnia+Herzegovina)*.

[63] Georgia law § 21-2-386(a)(2)(B) covers a similar situation:
"processing and scanning absentee ballots...While viewing or monitoring the process set forth in this paragraph, monitors and observers shall be prohibited from:... (vii) Communicating any information that they see while monitoring the processing and scanning of the absentee ballots, whether intentionally or inadvertently, about any ballot, vote, or selection to anyone other than an election official who needs such information to lawfully carry out his or her official duties."
Judge Boulee said broader wording would be needed to cover the entire period before polls close, since this "applies only ***during*** the viewing or monitoring of the absentee ballot opening and scanning process"
https://bradblog.com/Docs/8-20-21-order-in-coalition-v-raffensperger.pdf
Current Georgia law is at Title 21, Chapter 2, Article 10, Section 386((a)(2)(B)
https://advance.lexis.com/container?config=00JAAzZDgzNzU2ZC05MDA0LTRmMDItYjkzMS0xOGY3MjE3OWNlODIKAFBvZENhdGFsb2fcIFfJnJ2IC8XZi1AYM4Ne&crid=877c6772-0ad2-4ac5-831f-dbbcdca680d0
[64] It's not hard or costly for locksmiths to improve locks a great deal: https://www.youtube.com/watch?v=7JlgKCUqzA0 *Consumer Report*s rates locks on picking ease, though they do not describe their tests.
https://www.consumerreports.org/products/door-locks-34477/door-lock-33770/view2/

If candidates and parties do not trust locks and observers to keep storage safe, they can try to convince the legislature to adopt earlier processing or multi-partisan guards. Early processing need not tally votes. All it needs to do is create files of ballot images, with hash values, so no one has early results.

If there is no trust and no early processing, no audit can remove, or perhaps even reduce, doubts.

Levels of problems to expect

Election systems are complex, to track voter eligibility, ballots and summing all the tallies for thousands of races. Mistakes happen. Billions of dollars and controversial public policies are at stake, so mistakes need to be caught and corrected.

There were 67,000 computer breaches reported in the US in 2018-2020.[65] Large companies averaged 23 cybersecurity incidents per company in 2020: 15 per company caused by careless or negligent employees and contractors, 5 incidents per company from criminal insiders, and 3 from credential theft.[66] In election companies and offices, that level of errors and dishonesty can lead to inaccurate election results. Election errors are not as closely tracked as embezzlement, but there were at least 200 errors in election machines from 2002-2008, many of which happened repeatedly in different jurisdictions.[67] More errors have happened since then. Election vulnerabilities include:

A. Programming mistakes – Election computers are complex, updated every year. All big computer systems have many bugs.[68] Bugs can shift winners as much as hacks can, and many hacks look like bugs. So bugs need to be taken seriously. Recovering from bugs solves many hacks.

B. Undetected[69] and unfixed[70] backdoors and other hacks – These are pervasive, from multiple adversaries.

C. Insiders – Staff at election system manufacturers, and their suppliers, and at election offices can cause harm when their credentials are phished, or if any of them can be bribed or blackmailed by people who want a contract or land use approval or other goal. Vulnerability to insider fraud is shown by the 13,500 arrests[71] for embezzlement per year in US businesses. 85% involved a

---

[65] 67,000 reports of computer breaches: https://www.securityweek.com/deep-analysis-more-60000-breach-reports-over-three-years Most were reported by the hackers. Only 13% were reported through official channels. These were not reported at election companies or local governments. If hackers don't reveal election theft, we won't know.

[66] *2020 Cost of **Insider Threats** Global Report*. Sponsored by ObserveIT and IBM https://go.proofpoint.com/wp-ponemon-itm-cost-of-insider-threats.html

[67] Brennan Center list of voting machine errors,"brennancenter.org/sites/default/files/2019-08/Report_Voting_Machine_Failures_Database-Solution.pdf In addition, Heritage has a database of 1,300 fraud from 1982-2000. Three quarters involved ineligible people voting, and the other quarter included fraud by election officials https://www.heritage.org/voterfraud/search Assessed by Brennan Center https://www.brennancenter.org/sites/default/files/2019-07/Report_HeritageAnalysis_Final.pdf

[68] **Prevalence of bugs:** Schneier, "Every piece of commercial software... has hundreds if not thousands of vulnerabilities, most of them undiscovered" https://www.google.com/books/edition/Data_and_Goliath_The_Hidden_Battles_to_C/MwF-BAAAQBAJ?hl=en&gbpv=1&dq=schneier&pg=PT135&printsec=frontcover

Examples of bugs and hacks in elections: http://votewell.net/a/hacks.htm

[69] **Undetected problems:** Kolbe on SolarWinds hack by Russia, *"Chinese, others, they've all built huge capabilities, they're well-resourced, well-staffed, [and] focused on doing exactly this. This is not a one-off, this is not something unusual... I guarantee you that there are other operations similar in size and scope, if not larger, that haven't been discovered."* https://news.harvard.edu/gazette/story/2020/12/harvard-cybersecurity-experts-discuss-russian-breach/

[70] **Unfixed problems:** Schneier. *"many network administrators won't go through the long, painful, and potentially expensive rebuilding process. They'll just hope for the best."* theguardian.com/commentisfree/2020/dec/23/cyber-attack-us-security-protocols

Amazon cybersecurity engineer says, "*Why would you care about cloud security? You don't have to bust your ass because you live in a small-town market where you know everybody and you're never going to be out of a job. A lot of companies that are headquartered in remote areas don't have particularly sophisticated IT teams."* https://logicmag.io/commons/inside-the-whale-an-interview-with-an-anonymous-amazonian/

[71] **Embezzlement arrests**: https://www.ojjdp.gov/ojstatbb/crime/ucr.asp?table_in=1

manager,[72] and most were long-time employees. In election companies and offices the same level of dishonesty can lead to election fraud instead of embezzlement.

D. Foreigners – Some want a policy outcome at one or more levels of government, or they want to destabilize by defeating incumbents. Foreigners can reach air-gapped machines through updates, and when memory devices go back and forth, to transfer results for public release on the internet. They can also act in person through local criminals.[73]

E. Criminals – Organized crime and gangs have an interest in choosing prosecutors, sheriffs or judges.

F. Picking locks – Picking polling place locks takes no sophistication, when machines sit unguarded the night before an election.

ImageCast Central

Canon DR-G1130
11" - 80/min, 4800/hr
14" - 64/min, 3840/hr
17" - 53/min, 3180/hr
20" - 45/min, 2700/hr

Canon DR-M160II
11" - 60/min, 3600/hr
14" - 47/min, 2820/hr
17" - 38/min, 2280/hr
20" - 33/min, 1980/hr

ImageCast Precinct Dominion

With ballot image capture enabled:
11" - 6/min, 360/hr
14" - 4.5/min, 280/hr
17" - 3.5/min, 230/hr
20" - 3.3/min, 198/hr

With ballot image capture disabled:
11" - 7.5/min, 450/hr
14" - 6/min, 360/hr
17" - 5/min, 300/hr
20" - 4/min, 240/hr

Do saved ballot images slow down scanning or overfill storage devices?

ES&S is the main company which does not automatically save ballot images. They say, "It is possible to turn off the ballot image storage feature. Disabling this feature equates to time savings in the *millisecond* range, which is unnoticeable to the user."[74]

Dominion central scanners only need about *a second* per ballot, regardless of saving images. Dominion precinct scanners take *3-5 seconds* per ballot extra to save images,[75] which is not enough to delay the next voter walking up to the scanner.

ES&S says for the DS200 precinct scanner, "If ballot images are captured and stored, the 4GB drive can store at least 12,000 images

		Memory Card Size			
Ballot Size (Single-Sided)	**Approx Ballot Image Size (KB)**	**2 GB**	**4 GB**	**8 GB**	**16 GB**
8.5" x 11"	250	6000	14000	30000	62000
8.5" x 14"	277	5400	12600	27000	55800
8.5" x 17"	312	4800	11200	24000	49600
8.5" x 20"	334	4500	10500	22400	46400
8.5" x 22"	357	4200	9800	21000	43400
Ballot Size (Double-Sided)					
8.5" x 11"	357	4200	9800	21000	43400
8.5" x 14"	454	3300	7700	16500	34100
8.5" x 17"	499	3000	7000	15000	31100
8.5" x 20"	555	2700	6300	13500	27900
8.5" x 22"	624	2400	5600	12000	24800

Dominion

Table 5-3: ImageCast® Evolution Ballot Image Capacity for Various Ballot Sizes and Memory Card Sizes

[72] **Who embezzles:** https://www.hiscox.com/documents/2018-Hiscox-Embezzlement-Study.pdf

[73] Besides access to storage by local criminals (and possibly dishonest insiders), foreigners are also a risk. Russia and North Korea are widely accused of **using criminals to act for them** abroad and there is no reason other countries cannot do the same:

https://ecfr.eu/article/commentary_security_concerns_russian_mafia_back_on_agenda7083/ "For Moscow, Russian-based criminal networks provide an unconventional asset in the geopolitical struggle with the West… At home and abroad, Russia's gangsters and spooks are often closely connected."

https://bpr.berkeley.edu/2019/12/16/gangs-and-gulags-how-vladimir-putin-utilizes-organized-crime-to-power-his-mafia-state/ "more and more crime networks tangentially linked to Russian actors have appeared all over Europe. Multiple politically convenient assassinations or assassination attempts have been made on anti-Russia figures by gang members who have been accused of being Russian assets... Russian government's implicit support of the vast array of organized crime means that the surge of activity across the European Union won't be going away any time soon."

https://globalinitiative.net/wp-content/uploads/2017/09/TGIATOC-Diplomats-and-Deceipt-DPRK-Report-1868-web.pdf "Much has been written about state-sponsored North Korean criminal activity in Asia and Europe."

https://www.dni.gov/files/documents/NIC_toc_foldout.pdf "North Korean entities maintain ties with crime networks to earn hard currency."

https://nautilus.org/napsnet/napsnet-policy-forum/the-north-korean-criminal-state-its-ties-to-organized-crime-and-the-possibility-of-wmd-proliferation/ "I compared North Korea under Kim Jong Il with Serbia under Milosevic, Romania under Ceausescu, and Panama under Noriega... Incidences of illicit activity have occurred in every continent and almost every DPRK Embassy in the world has been involved at one time or another."

[74] P.12 of Election Systems & Software, Llc Contract No. 071B7700120, Exhibit 2, Attachment 1.1 Voting System Hardware Technical Requirements https://www.michigan.gov/documents/sos/071B7700120_ESS_Exh_2_to_Sch_A_Tec_Req_555360_7.pdf

[75] Dominion Voting Systems Contract No. 071B7700117 Exhibit 2, Attachment 1.1 Voting System Hardware Technical Requirements michigan.gov/documents/sos/071B7700117_Dominion_Exhibit_2_to_Sch_A_Tech_Req_555357_7.pdf

(both sides of each ballot sheet), along with the cast vote records, the election definition, and audit log records…" and for central scanners, "The DS450 contains a 1-terabyte internal hard drive that stores all of the ballot images and records for the entire election, no matter the size or complexity of the election. Contractor estimates it will take an entire man year of continuous scanning to fill the hard drive."[76]

Dominion says 5,000-14,000 images fit on a 4GB card and 24,000-62,000 images fit on a 16GB card, depending on ballot size, as shown in table.[77]`

COSTS

Audit costs for 100,000 ballots are in Table 2, with a link to a spreadsheet where the number of ballots and assumptions can be changed. below. Table 2 gives examples of contracts for analysis of existing ballot images in Clackamas and Montrose Counties. Table 2 gives the cost in New York to buy equipment for this approach and use it for 12 years, plus three cents per ballot for local staff time to run the equipment. Table 2 also gives the cost to contract for this approach in Clackamas County. Buying equipment has about the same cost as paying outsiders to do BIAs when the equipment handles two or more elections per year. Buying equipment would cost more per election in the many states which do not have two elections every year. Having a local group adapt open source BIA software is also an option (Elections Transparency Project, 2022; Free and Fair, 2017).

The cost of an RLA for 1-2 contests depends on the margin of victory in those contests. Table 2 shows an example where 300 ballots are enough.

The cost of RLAs for all contests depends on how many ballots were cast in the closest contests in each area. Detailed analysis of Hennepin County (Minneapolis) and Allegheny County (Pittsburgh) shows a range of 1.8% to 7% of all ballots needed to cover all local contests with *ballot comparison RLAs*. *Batch comparison RLAs* cost more and are covered with other batch audits below. Rhode Island estimated 30 clock minutes (60 staff minutes for a team of two) are enough to collect 10 boxes from storage and return them in a storage area of 9,000 ballots (RI RLA Working Group, 2019). 100,000 ballots need 11 times as much space, so distances will be 3.3 times as far ($\sqrt{11}$), and 100 minutes (200 person-minutes) per 10 boxes are plausible. Offices will find that shelves let them get to all boxes easily. Rhode Island also estimated time to pull samples from boxes and examine them, which is included in the minutes per batch shown. Arlo and Minerva software is available free (VotingWorks, 2022; Zagórski et al., 2021). Maryland's pilot RLA paid $2,600 for consultation from a statistics professor (Maryland State Board of Elections, 2016).

Costs assume $30/hour, based on $15/hour for staff time working directly on the ballots, plus 100% overhead for recruiting, training and supervision. The spreadsheet shows detail. Other pay rates will have proportional costs. Background checks on staff are rarely done for temporary election workers and would add to the costs shown.

RLAs cost much more when a jurisdiction-wide contest is so close it needs a 100% hand count of the whole jurisdiction. Each jurisdiction averages 9 jurisdiction-wide contests, which include: 5 state-wide (Ballotpedia, 2022), 3 county-wide (NACo, 2019), and 1 congressional. Among congressional races in 2020, one (0.23%) was so close that an RLA would need a 100% hand count (Iowa-2). If each of the 9 contests has 0.23% chance of needing a full 100% hand count, the overall chance that at least one of them will need a full hand count is 2%. Hand counting one contest would cost $25,000, a tenth of the $250,000 for a full hand count of 10 contests at the bottom of Table 2. A 2% chance of spending $25,000 means an average cost of $500. This does not cover the chances or cost of a full hand count needed when auditors find problems.

Ballot comparison RLAs need to match ballots and computer records in the same order. About half of ballots are scanned in polling places and fall into ballot boxes in random order. Therefore Table 2 includes the cost of rescanning half the ballots centrally, to have ballots and records in the same order, based on Rhode Island's experience. An office with more or fewer ballots scanned in polling places can adjust costs of rescanning in Table 2.

Quality control samples involve comparing a random sample of paper ballots to their electronic records. Random selection and comparison take 3.1 minutes per sampled ballot by teams of two, so 6.2 staff minutes per sampled ballot. For this one-to-one matching, the ballots and records need to be in the same order as each other, entailing the same rescanning cost for ballots scanned in polling places as described above for RLAs.

---

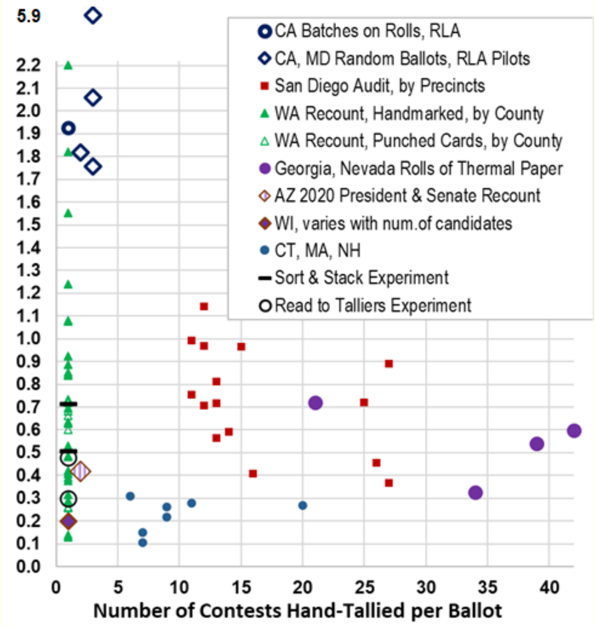
[76] P.27 of Election Systems & Software, Llc Contract No. 071B7700120, Exhibit 2, Attachment 1.1 Voting System Hardware Technical Requirements https://www.michigan.gov/documents/sos/071B7700120_ESS_Exh_2_to_Sch_A_Tec_Req_555360_7.pdf

[77] page 77 of 2.02 - Democracy Suite® System Overview, Version: 5.11-CO::7, June 5, 2019 sos.state.co.us/pubs/elections/VotingSystems/DVS-DemocracySuite511/documentation/2-02-SystemOverview-5-11-CO.pdf

Batch percentage audits and batch-comparison RLAs require a large number of batches, usually all, to audit the close contests which appear in almost every local election. Table 2 shows a hypothetical sample of 20% of batches, but this is rarely enough for reliable audits of small contests. To select the sample, batch audits need an inventory of all batches with candidate totals for each batch. This is the same sampling frame that ballot comparison RLAs use.

Hand count timing has been measured in 10 studies. The time was usually 0.1 to 1.2 minutes of staff time per vote, for teams of 3-4 staff (Figure 6). Table 2 uses an estimate below the middle of this range: half a minute of staff time (7.5 seconds for a team of 4) to count each vote and resolve discrepancies. Table 2 assumes 10 votes per ballot times half a minute each, so 5 staff minutes per ballot. Using volunteers would save their pay and cost about 40% as much; volunteers still need training and supervision. The recruitment process for staff or volunteers need to include all political groups, since the public will only trust bipartisan teams, or multi-partisan if possible.

Figure 6
Hand Tallies Usually Need 1/3 to 1 Minute per Vote
Staff-minutes per vote, excludes training, setup, administration



Total staff-minutes = ballots X contests/ballot X minutes/vote
Cost = (pay/hour + cost-of-training&admin/hour) X staff-minutes/60
Clock-minutes needed = staff-minutes / number-of-staff
Source: commons.wikimedia.org/wiki/File:Time_for_hand_counts.png

Table 2. Costs of Ballot Image Audits (BIAs), Risk-limiting Audits (RLAs), Quality Control and Hand Counts For a place with 100,000 ballots. Other sizes roughly proportional. Costs can be 30% higher or lower, especially if problems are found.	Number of Ballot Sheets Examined	Fixed Costs per Jurisdiction per Election	***Dollars*** or **Staff Minutes** per Sheet	Number of Batches	Staff Minutes per Batch	Total Cost
BIA, existing images (Clackamas OR, 1 election)	100,000		***$0.17***			**$17,000**
BIA, existing images (Montrose CO, 1 election)	100,000	$12,200	***$0.25***			**$37,200**
BIA, rescan (NY 12 yrs, 24 elections)	100,000	$9,000	***$0.265***			**$35,500**
BIA, rescan (Clackamas OR, 1 election)	100,000		***$0.26***			**$26,000**
Ballot comparison RLA of 1-2 contests, 300 ballots	300			226	34	$3,453
Ballot comparison RLA of all contests, 1.8% of ballots	1,800			486	46	$11,216
Ballot comparison RLA of all contests, 7% of ballots	7,000			500	99	$24,833
2% chance of 100% hand count of one contest						$500
Rescan in-person batches to have in order for sampling	50,000		***$0.03***			$1,500
Statistics consultant, using free software		$2,600				$2,600
Ballot comparison RLA of 1-2 contests, 300 ballots, total						**$7,553**
Ballot comparison RLA of all contests, 1.8% of ballots, total						**$15,816**
Ballot comparison RLA of all contests, 7% of ballots, total						**$29,433**
RLA software & training, optional		$0-168,000				$0-168,000
RLA software & 1 session of each training & 1 day of each support	100,000	$5,500				$5,500
Quality control (QC) while scanning for sample	600		**6.2**			$1,860
Rescan in-person batches to get in order for sampling	50,000		***$0.03***			$1,500
Total QC cost						**$3,360**
Batch RLA or percentage audit, 1 contest 1% of batches	1,000		**0.5**	5	127	**$318**
6 contests 5% of batches	5,000		**3.0**	25	627	**$7,838**
All contests 1% of batches	1,000		**5.0**	5	1,027	**$2,568**
All contests 5% of batches	5,000		**5.0**	25	1,027	**$12,838**
All contests 20% of batches	20,000		**5.0**	100	1,027	**$51,350**
All contests 20% of batches and QC	20,000		**5.0**	100	1,027	**$54,710**
100% hand count	100,000		**5.0**			**$250,000**
Counts by volunteers. Paid trainers & supervisors.	100,000		**5.0**			**$103,423**

Source: votewell.net/tallies.xlsx. Notes: RLA with quality control only pays to rescan once. “All contests” will cost more when there are more than 10 contests per ballot.