QA HEALTH CHECK · AUDIT REPORT

● 14 FINDINGS · 4 HIGH PRIORITY

Acme Ledger · QA Health Check

Period	17 Aug 2026 – 28 Aug 2026 (2 weeks)
Scope	Web app and Public API; release process; Xray; CI; UAT; environments
Access	Read-only: Jira, Confluence, Xray, GitHub, Sentry, staging
Interviews	CTO, 2 engineering leads, product owner, Finance, Support
Data analysed	1,240 Xray cases, 12 releases, 488 E2E tests, 54 open bugs
Prepared by	Tomasz Kordiak (QA lead)
Report date	28 Aug 2026

Contents

1. Executive summary
2. Snapshot
3. Scores by area
4. Findings
5. Defects found during the audit
6. Observations
7. Risk map
8. Test management inventory
9. Automation inventory
10. Release history
11. Defect backlog
12. Interviews
13. 30 / 60 / 90-day plan
14. KPIs to track
15. What you keep

Executive summary

Acme ships every two weeks with a capable team and a solid UI test suite, but nobody can say with confidence what was tested before a release. Tests exist; the process around them doesn't: no exit criteria, an outdated test management tool, UAT in Slack, no API tests and no release report. 8 of the last 12 releases needed a hotfix within a week (11 hotfixes in total).

The four changes with the best return are a release checklist with exit criteria, an API smoke suite in CI, recorded UAT, and a cleaned-up Xray linked to tickets. Together they take about 6 weeks with the current team plus part-time QA support.

> **Biggest risk found**
>
> Users with the read-only Accountant role can export every customer's contact details as CSV (F-01, ACME-471). That is personal data beyond what the role needs.

Snapshot

1,240
Test cases in Xray

65%
Not run in 6 months

65%
Not linked to a ticket

488
E2E tests (Playwright)

0
API tests

8 / 12
Releases with a hotfix

Scores by area

AREA	SCORE	WHY
Release process	2/5	No written exit criteria; go/no-go decided in a chat thread the evening before.
Requirements and acceptance criteria	2/5	31% of stories in the last 3 releases had no acceptance criteria.
Test management (Xray)	2/5	1,240 cases; 38% not run in 6 months; 41% not linked to any ticket.
Coverage of risky areas	3/5	Payments and auth well covered; data retention, exports and roles barely.
Test automation (UI)	3/5	488 Playwright tests, healthy; 6% of runs have flaky failures.
Test automation (API)	1/5	No API tests; integrations rely on UI tests and customers.
CI and quality gates	3/5	E2E runs on merge; failures don't block the release branch.
Environments and test data	2/5	Staging shared by 3 teams; data reset by hand, about monthly.
Defect management	4/5	Clear templates and priorities; 23 bugs older than 90 days with no owner.
UAT with the business	2/5	Finance tests from screenshots in Slack; results not recorded.

AREA	SCORE	WHY
Non-functional testing	1/5	No load, security or accessibility testing in the last 12 months.
Reporting to management	1/5	No release report; management asks engineers what was tested.

Average 2.2 / 5. Scale: 1 = missing, 3 = works but depends on individuals, 5 = documented, measured and routine.

Findings 14 findings

ID	FINDING	PRIORITY	AREA	EFFORT
F–01	Read-only role can export all customer contact data	High	Security / GDPR	S
F–02	No release exit criteria or release report	High	Release process	S
F–03	UAT results are not recorded	High	UAT	S
F–04	No API tests; integrations untested before release	High	Automation	M
F–05	Test management has drifted from the product	Medium	Xray	M
F–06	Data retention and deletion untested	Medium	Coverage	S
F–07	Shared staging with manual data resets	Medium	Environments	M
F–08	E2E failures don't block the release branch	Medium	CI	S
F–09	31% of stories without acceptance criteria	Medium	Requirements	S
F–10	No load, security or accessibility testing	Medium	Non-functional	L
F–11	Flaky E2E tests are retried, not fixed	Low	Automation	S
F–12	23 bugs older than 90 days without an owner	Low	Defects	S
F–13	Bug reports from support lack reproduction steps	Low	Defects	S
F–14	Old test folders from 2024 teams still active	Low	Xray	S

Effort: S = under 2 days, M = up to 2 weeks, L = more. Details for the high-priority findings:

High F-01 **Read-only role can export all customer contact data**

Evidence	As `Accountant`, **Customers → Export CSV** downloads 312 customers with names, emails and phone numbers. The role description says 'invoices and reports only'.
Impact	Personal data available to more people than needed (GDPR data minimisation); easy to leak by email.
Recommendation	Restrict the export to Admin and Finance; add a role-permission matrix test.

High F-02 **No release exit criteria or release report**

Evidence	Last 12 releases: go/no-go discussed in Slack; no record of what was tested; 11 hotfixes within 7 days.
Recommendation	One-page checklist with exit criteria; a release report from Xray for each release.

High F-03 **UAT results are not recorded**

Evidence	Finance tests from screenshots posted in `#release`; 3 of the last 4 UATs have no record of which flows were checked.
Recommendation	Plain-English UAT cases in Xray, assigned per tester; a screenshot per result.

High F-04 **No API tests**

Evidence	69 documented API operations, 0 automated tests. 4 of the last 12 hotfixes were API regressions reported by customers.
Recommendation	Run the 126 smoke checks written during this audit in CI on every staging deploy, then extend to all resources.

Defects found during the audit 8 reported in Jira

KEY	SUMMARY	SEVERITY	STATUS	COMPONENT	FOUND BY	OWNER	CREA
`ACME-471`	Accountant role can export every customer's contact details as CSV	High	Open	Roles	`Exploratory`	Backend	19 Au
`ACME-472`	Invoice PDF shows the old company address after a settings change	Medium	Open	PDF	`Exploratory`	Backend	19 Au

KEY	SUMMARY	SEVERITY	STATUS	COMPONENT	FOUND BY	OWNER	CRE
`ACME-473`	`GET /customers?country=xx` returns 500 instead of 400	Medium	Open	Customers API	`API smoke`	Backend	20 A
`ACME-474`	Webhook delivery log shows times in UTC without a label	Low	Open	Webhooks	`Exploratory`	Frontend	20 A
`ACME-475`	Session not invalidated after password change	Medium	Open	Auth	`Exploratory`	Backend	21 Au
`ACME-476`	Report CSV uses comma decimal separator for English locale	Low	Open	Reports	`API smoke`	Backend	21 Au
`ACME-477`	Portal payment page has no visible focus indicator	Low	Open	Customer portal	`Exploratory`	Frontend	24 A
`ACME-478`	OpenAPI spec lists `DELETE /customers/{id}` which returns 405	Low	Open	API docs	`API smoke`	Backend	25 A

Observations

ID	AREA	OBSERVATION	IMPACT	SUGGESTED ACTION
`OBS-01`	Team	Developers write good unit tests (78% line coverage on the backend).	Positive	Keep; the gap is above unit level.
`OBS-02`	Automation	The Playwright suite is well structured and fast (11 min on 4 shards).	Positive	Build on it; no rewrite needed.
`OBS-03`	Process	Hotfixes skip review on Fridays to 'get them out before the weekend'.	Risk	Same checklist for hotfixes; smoke test before deploy.
`OBS-04`	Tools	Xray and a Google Sheet are both used to track UAT; they disagree in 14 places.	Effort	Xray only (finding F-03).

ID	AREA	OBSERVATION	IMPACT	SUGGESTED ACTION
`OBS-05`	Monitoring	Production errors are visible in Sentry but nobody reviews them before a release.	Risk	Add 'new errors since last release' to the release checklist.
`OBS-06`	Support	Support files 6–10 bugs a month; 40% are duplicates or questions.	Effort	Weekly 20-minute triage with QA.
`OBS-07`	Docs	Confluence has a test strategy from 2023 that no one follows.	Low	Replace with the one-page checklist.

Risk map

PRODUCT AREA	BUSINESS IMPACT	CHANGE FREQUENCY	TEST COVERAGE	RISK
Webhooks and API	4	3	1	60
Data retention and deletion	5	2	1	50
Roles and permissions	4	3	2	48
Customer portal	4	4	3	48
Invoice totals and VAT	5	4	4	40
Reports and exports	3	3	2	36
Recurring invoices	4	2	2	32
Payments and refunds	5	3	4	30
Search	2	2	2	16
Settings	2	2	3	12

Each scored 1–5. Risk = impact × change frequency × (6 − coverage). Highest-risk areas get tests first.

Test management inventory

XRAY FOLDER	CASES	RUN IN 6 MONTHS	LINKED TO TICKET	AUTOMATED	LAST RUN
Invoices	214	61	58	52	2026-08-12

XRAY FOLDER	CASES	RUN IN 6 MONTHS	LINKED TO TICKET	AUTOMATED	LAST RUN
Payments	168	74	66	41	2026-08-12
Customers	96	52	47	30	2026-07-30
Credit notes	54	18	20	9	2026-06-03
Recurring	72	21	25	12	2026-05-21
Reminders	48	30	28	14	2026-07-30
Reports	88	31	36	18	2026-06-17
Users and roles	64	40	41	24	2026-08-12
Customer portal	92	70	61	38	2026-08-12
Settings	58	22	19	8	2026-04-08
Integrations	41	9	12	4	2026-02-25
Team A (2024)	82	0	12	0	2025-11-19
Team B (2024)	67	0	9	0	2025-10-02
Old regression	71	0	4	0	2025-06-30
Misc / unsorted	25	0	2	0	2025-09-15
Total	**1240**	**428**	**440**	**250**	

Highlighted folders belong to 2024 team structures and haven't been run since; archive candidates.

Automation inventory

SUITE	TESTS	RUNTIME	FLAKY RUNS	RUNS ON	OWNER
Backend unit tests (Jest)	2,940	3 min	0%	Every PR	Backend
Frontend unit tests (Vitest)	1,210	2 min	0%	Every PR	Frontend
E2E (Playwright, Chromium + WebKit)	488	11 min	6%	Merge to main	Frontend
API tests	0	–	–	–	Nobody
Load tests	0	–	–	–	Nobody

SUITE	TESTS	RUNTIME	FLAKY RUNS	RUNS ON	OWNER
Visual regression	0	–	–	–	Nobody

Release history

RELEASE	DATE	TICKETS	HOTFIXES ≤ 7 DAYS	CUSTOMER-FOUND DEFECTS
3.9	2026-03-04	31	1	3
3.10	2026-03-18	27	–	1
3.11	2026-04-01	35	2	4
3.12	2026-04-15	22	–	2
3.13	2026-04-29	29	1	2
3.14	2026-05-13	40	1	5
3.15	2026-05-27	26	–	1
3.16	2026-06-10	33	1	3
3.17	2026-06-24	38	2	6
3.18	2026-07-08	24	–	2
3.19	2026-07-22	30	1	3
4.0	2026-08-05	46	2	7

Over 12 releases: 11 hotfixes and 39 defects found by customers.

Defect backlog

SEVERITY	< 30 DAYS	30–90 DAYS	90–180 DAYS	> 180 DAYS	TOTAL
Critical	–	–	–	–	**0**
High	2	1	1	–	**4**
Medium	9	7	6	8	**30**

SEVERITY	< 30 DAYS	30–90 DAYS	90–180 DAYS	> 180 DAYS	TOTAL
Low	14	11	12	15	**52**

86 open bugs; 42 older than 90 days, 23 of them with no owner.

Interviews

WHO	KEY POINT
CTO	“I find out what was tested when a customer tells me what wasn't.”
Engineering lead (backend)	Wants API tests but nobody owns them; unit tests are strong.
Engineering lead (frontend)	Playwright suite is trusted; flaky tests are annoying but known.
Product owner	Acceptance criteria are written 'when there's time'; UAT is the safety net.
Finance (UAT)	Doesn't know which flows to test; uses last release's Slack thread as a checklist.
Support lead	Most escalations are permissions and exports; bugs are hard to reproduce.

30 / 60 / 90-day plan 17 tasks

WHEN	TASK	OWNER	EFFORT	DONE WHEN
Days 1–30	Fix ACME-471 (export permission)	Backend	1 d	Fix released and tested for all 5 roles
Days 1–30	One-page release checklist with exit criteria	QA + Eng lead	1 d	Used for the next 2 releases
Days 1–30	Release report per release from Xray	QA	2 d	Report sent to management for 4.1 and 4.2
Days 1–30	API smoke suite in CI (126 checks written in the audit)	QA	2 d	Runs on every staging deploy
Days 1–30	Triage the 23 old bugs	PO + QA	1 h	Each closed, scheduled or owned

WHEN	TASK	OWNER	EFFORT	DONE WHEN
Days 1–30	Acceptance criteria required in Jira workflow	PO	1 h	No story enters sprint without them
Days 31–60	Restructure Xray by product area; archive 4 old folders	QA (with agent)	3 d	Every 4.x ticket has linked cases
Days 31–60	Plain-English UAT cases for Finance, Support, Controlling	QA	3 d	UAT for 4.1 recorded in Xray
Days 31–60	Seed script and nightly staging reset	Platform	3 d	Data reset without manual work
Days 31–60	E2E failures block the release branch	Eng lead	1 h	Branch protection rule active
Days 31–60	Data retention tests (API + E2E)	QA	2 d	Tests in CI before retention ships
Days 31–60	Extend API suite to all resources	QA	4 d	≥ 250 checks, ≥ 90% of operations
Days 61–90	Fix the 2 known flaky tests	QA	1 d	< 1% flaky runs over 2 weeks
Days 61–90	First load test at expected peak	QA	3 d	Capacity known before January
Days 61–90	Penetration test of web app and API	QA / security	5 d	Report and retest
Days 61–90	Accessibility audit of key screens	QA	3 d	WCAG 2.2 AA status known
Days 61–90	Weekly support bug triage	QA + Support	20 min/week	Duplicate rate < 15%

KPIs to track

METRIC	TODAY	TARGET IN 90 DAYS
Hotfixes within 7 days of release	11 in 12 releases	≤ 1 per release
Defects found by customers per release	3.25 (average)	≤ 1
Stories with acceptance criteria	69%	100%

METRIC	TODAY	TARGET IN 90 DAYS
Xray cases run in the last 6 months	35%	≥ 90%
Cases linked to a ticket	35%	≥ 95%
API operations covered by tests	0%	≥ 90%
Flaky E2E runs	6%	< 1%
Time to produce a release report	Not produced	≤ 5 minutes

What you keep

This report and the raw data behind every number; 126 API smoke checks in your repository (`/qa/api-smoke`); a draft release checklist and UAT case template in Confluence; the risk map as a spreadsheet.

Sample report. Acme Ledger, its data, people and defects are fictional. Your report shows your product, your tickets and your tools.

More sample reports · Book a 15-minute call