

---

## TESTING RECORD

# IGCE Builder CR Testing Record

---

*Four waves of testing across April 2026 shipped 35 patches to the Cost-Reimbursement IGCE skill, including a critical fee-on-ODCs correctness fix that prevented \$1.1M of silent over-fee across the regression scenarios.*

---

### Skill

igce-builder-cr  
[github.com/1102tools/federal-contracting-skills](https://github.com/1102tools/federal-contracting-skills)

### Date

April 2026  
[1102tools.com](https://1102tools.com)

# Part 1: For Federal Acquisition Users

---

## The bottom line

Four rounds of testing for the Cost-Reimbursement skill across April 2026: Wave 1 inherited the full patch set from FFP Wave 5 and LH/T&M Wave 2 without direct CR scenarios. Wave 2 ran six scenarios on Claude Opus 4.7 across two rounds (three lazy-prompt, three detailed-prompt), shipping 14 patches plus 11 universal-principle patches ported across all 3 IGCE skills (Wave 3). Wave 4 ran 4 cold tests targeting untested territory (CO-supplied DCAA rates, FCCM layer, pass-through ODCs with fee implications, CPFF Completion vs Term Form), shipping 4 universal patches including one critical correctness fix. 4 regression agents validated all patches hold with zero regressions. The skill now produces auditable CPFF (Completion or Term Form), CPAF, and CPIF workbooks end-to-end with layered cost pool buildup, Fee-Bearing vs Non-Fee Cost separation, DCAA-rate override support, FCCM treatment, fee structure analysis, and ai-boundaries-compliant narrative.

- **Wave 1** (inherited): six universal patches from FFP Wave 5 and LH/T&M Wave 2 applied without direct CR scenarios. ai-boundaries v2 gate, pre-flight MCP check, Step 0 two-stage validation gate (with fee type as a Stage B parameter), DoD installation to GSA per diem crosswalk, multi-destination travel sheet, CLI recalc fallback, CALC+ query optimizations, FY rollover guidance.
- **Wave 2** (six scenarios validated, Claude Code CLI, Opus 4.7): three lazy-prompt scenarios plus three detailed-prompt scenarios run with realistic user framing to exercise decomposition, parameter prompting, and fee-type selection across all three fee types. Lazy round: 22 findings, 14 patches shipped. Detailed round: 31 additional findings, 11 triaged as universal patches shipped to all 3 IGCE skills (see Wave 3).
- **Wave 3** (universal patches derived from CR Wave 2 detailed round): 11 universal-principle patches shipped identically to FFP, LH/T&M, and CR. Inherited across all 3 skills; not re-tested per-skill.
- **Wave 4** (4 cold tests on post-Wave-3 skill, 4 patches shipped, 4 regression tests passed): targeted untested territory from "What has NOT been tested" list. Fee-on-ODCs critical correctness fix (CPAF test showed \$959K over-fee avoided across 5 years; CPIF test showed \$120K avoided across 4 years; CPFF test showed \$29K avoided across 3 years). CPFF Completion vs Term Form prescription with FAR 16.306(d)(1) vs (d)(2) citation and decision heuristic. CO-supplied DCAA rates override rule (use FPRA rates instead of 32/80/12 defaults when CO supplies). FCCM (FAR 31.205-10 / CAS 414) as distinct cost pool layer between G&A

and Total Estimated Cost. Regression agents validated all 4 patches held across CPFF / CPAF / CPIF with FPRA rates, FCCM > 0, and pass-through ODCs.

## Scenarios tested and how reliably they work

Scenario	Fee type	Prompt style	Result
CPFF biomedical research at NIH Bethesda (PhD biomedical scientists, no travel, base + 3 OYs)	CPFF	Lazy: "price an NIH research contract, 4 PhDs, base plus options"	Reliable after Wave 2 patches. Medical Scientist SOC (19-1042) added. Cost pool buildup at research-lab defaults. 85% assumed earned gated.
CPAF managed services at civilian agency (10-person team, CALC+ rate validation, base + 2 OYs)	CPAF	Lazy: "build a CR IGCE for a managed services contract with award fee"	Reliable after Wave 2 patches. 3-scenario fee view (base only, target with 85% earned, ceiling with full pool) shown in Summary.
CPIF Oak Ridge DOE engineering (asymmetric share ratio, bound-crossing variance, 6 LCATs, base + 4 OYs)	CPIF	Lazy: "CPIF IGCE at Oak Ridge, 80/20 over and 50/50 under, complex work"	Reliable after Wave 2 patches. Asymmetric share ratios split into contractor_share_over and contractor_share_under. ±25% bound-crossing variance documented.
CPFF AFRL BAA (Dayton)	CPFF 8%	Detailed prompt: "cpff for a BAA with AFRL out of Wright-Patt. R&D advanced materials characterization..."	Valid workbook, \$5.61M 5-yr. Workflow A+ clean. FY rollover patch worked. Wright-Patt → Dayton crosswalk confirmed. 9 findings.
CPAF HHS Claims (DC+Baltimore)	CPAF 3%+7%	Detailed prompt: "cpaf igce for HHS O&M claims processing..."	Valid workbook, \$24.6M 5-yr. Multi-metro labor split (8 DC / 4 Baltimore). 3-scenario fee view confirmed. 10 findings.
CPIF Sandia RF (Albuquerque)	CPIF 7.5% target, 80/20 over,	Detailed prompt: "cpif for Sandia, 8 RF hardware engineers..."	Valid workbook, \$12.27M 3-yr. Asymmetric share ratio confirmed. DOE crosswalk confirmed. 12 findings.

Scenario	Fee type	Prompt style	Result
	70/30 under		

## Manual-verification checklist

Scan every CR IGCE output for these before using in a contract file:

- 1. Fee type declared in Summary row 5.** CPFF, CPAF, or CPIF label is mandatory. Fee math downstream cascades from this cell.
- 2. Cost pool buildup is layered, not collapsed.** Direct Labor → Fringe → Labor+Fringe → Overhead → Subtotal → G&A → Total Cost → Fee → Total Price. Methodology must explain each layer.
- 3. Statutory fee caps enforced.** R&D fee cannot exceed 15% of estimated cost per 10 USC 3322(a). Non-R&D practical ceiling is 10%. If fee exceeds these, flag and reduce.
- 4. CPIF share ratios asymmetric by default.** Most real CPIF agreements have different over and under share ratios (e.g., 80/20 over, 50/50 under). The assumption block exposes both as separate variables.
- 5. CPIF bound-crossing documented.** When the overrun variance is wide enough that fee hits the min bound, or underrun variance hits the max bound, Methodology must explicitly say "share ratio stops applying at this cost level."
- 6. CPAF shows 3 fee scenarios, not 1.** Summary must show: (a) base fee only (worst earned), (b) base + 85% pool (target), (c) base + full pool (ceiling). Single-point assumption hides range from CO.
- 7. ai-boundaries gate held.** First response to "is this CR rate reasonable" must emit the refusal template, not a determination.

## What the skill does not do

- It does not produce FFP or LH/T&M estimates.** Use IGCE Builder FFP or IGCE Builder LH/T&M.
- It does not substitute for a contracting officer's price reasonableness determination.** IGCE is an estimate; the CO makes the determination per FAR 15.404.

- **It does not enforce agency-specific fee policies.** Some agencies cap CR fee below 10 USC statutory limits; skill flags the statutory cap only.
- **It does not produce a formal DCAA-compliant cost proposal review.** IGCE is the government-side estimate; DCAA audits the contractor's proposed rates separately.
- **It has not been tested on:** BAA cooperative agreements under FAR 35.016 with non-standard fee structures, Termination for Convenience cost estimation, CR-to-FFP conversion modeling, or OCONUS CR builds.

---

## Part 2: For Developers and Technical Reviewers

---

### Testing methodology

#### Scenarios

Three scenarios designed to exercise distinct CR mechanics across fee types:

- **S1 (CPFF, biomedical research):** NIH Bethesda, 4 PhD medical scientists, no travel, base + 3 OYs. Lazy prompt: "price an NIH research contract, 4 PhDs, base plus options." Exercises Medical Scientist SOC mapping (19-1042 vs 19-1099 Life Scientists All Other), NIH-domain cost pool defaults, CPFF 8% fixed fee, statutory 15% R&D cap awareness, no-travel Sheet 5 handling.
- **S2 (CPAF, managed services):** Civilian agency, 10-person team, 4 travel destinations, CALC+ rate validation, base + 2 OYs. Lazy prompt: "build a CR IGCE for a managed services contract with award fee." Exercises CPAF 3-scenario fee view (base/target/ceiling), assumed-earned gating, civilian-agency cost pool defaults, multi-destination Sheet 5 parameterization.
- **S3 (CPIF, Oak Ridge DOE engineering):** 6 LCATs (Mechanical, Electrical, Chemical, Nuclear Engineers, PM, Admin), asymmetric 80/20 over and 50/50 under share ratios,  $\pm 10\%$  baseline and  $\pm 25\%$  bound-crossing variance, base + 4 OYs. Lazy prompt: "CPIF IGCE at Oak Ridge, 80/20 over and 50/50 under, complex work." Exercises CPIF asymmetric share ratio support, bound-crossing documentation, DOE M&O cost pool defaults, DOE lab per diem crosswalk (Oak Ridge → Knoxville TN).

Each scenario had a 14-point binary assertion matrix covering skill activation, fee type selection, cost pool layering, rate validation, FAR citation completeness, workbook structural integrity, methodology completeness, and lazy-prompt recovery (did the skill prompt for missing inputs rather than guess).

## Environment

- Claude Code CLI, fresh conversation per scenario, Opus 4.7
- Local ~/ .claude/skills/igce-builder-cr/SKILL.md post-Wave 1 inheritance
- All three scenarios completed in a single worker pass without "continue" (post-inheritance skill was slim enough)

## Grading

Grader read worker's final response plus produced xlsx. Workers not coached. Assertions graded binary pass/fail. Worker self-critiques incorporated when corroborated by direct observation.

## Wave 1 (inherited, not directly tested on CR)

All universal patches derived from FFP Wave 5 and LH/T&M Wave 2 testing applied to CR at Wave 1:

- **ai-boundaries positioning (v2 gate):** Workflow B Step 0 token-scan + verbatim refusal template. Skill does not originate "fair and reasonable" determinations, price reasonableness memos, or negotiation recommendations.
- **Pre-flight MCP dependency check:** validates bls-oews, gsa-calc, gsa-perdiem tools and API keys before any workflow runs.
- **Step 0 two-stage validation gate:** Stage A (decomposition) + Stage B (build parameters including fee type CPFF/CPAF/CPIF) as separate AskUserQuestion calls. Skip for Workflow A with structured inputs.
- **DoD installation to GSA per diem crosswalk:** 15-row table mapping military installations to GSA civilian localities.
- **Multi-destination travel sheet:** Sheet 5 parameterized for M destinations, Sheet 1 Travel SUMs across blocks.
- **CLI recalc fallback:** Python expected-total check when LibreOffice recalc.py unavailable.
- **Step 9 environment fork:** delivery path varies by environment (claude.ai / Claude Code CLI / macOS Desktop with Numbers).

- **CALC+ query optimizations:** keyword\_search to igce\_benchmark for stats-only; tier-matched keywords to avoid false divergence flags.
- **FY rollover guidance:** if contract PoP start within 6 months of next FY, query both and document refresh on publication.
- **Raw Data sheet granularity:** summary tables with query parameters, not raw JSON dumps.

Wave 2 results (lazy-prompt validated)

Scenario	Score	Fee type
S1 CPFF biomedical research	14/14	CPFF
S2 CPAF managed services	14/14	CPAF
S3 CPIF Oak Ridge engineering	14/14	CPIF
Total	42/42 (100%)	—

Wave 2 findings: 22 surfaced, 14 patched

Universal patches (horizontal, ported to FFP and LH/T&M)

1. **Installation to GSA locality crosswalk expanded with 6 DOE labs.** Oak Ridge/Y-12 to Knoxville, LANL to Santa Fe, Hanford/PNNL to Richland, Sandia to Albuquerque, LLNL to Oakland-Fremont, INL to Idaho Falls. Prevents empty per diem lookups for DOE R&D scope.
2. **BLS MSA URL fallback for metros outside list\_common\_metros.** Worker hit this in S1 when the NIH Bethesda metro wasn't in the common-metros list. Skill now directs worker to resolve MSA code via [https://www.bls.gov/oes/current/msa\\_def.htm](https://www.bls.gov/oes/current/msa_def.htm) rather than silently falling back to state wages.
3. **Workflow A ambiguous-input rule.** Lazy prompts like "4 PhDs" without discipline triggered worker guessing. Skill now requires AskUserQuestion for ambiguous required inputs before pulling data.
4. **Step 9 env fork with macOS Excel/Numbers branch.** Claude Code CLI with Excel or Numbers installed can skip the Python-side expected-total check; open triggers recalc via system handler.



5. **BLS wage-cap 10% proximity rule.** When chosen percentile lands within 10% of the \$239,200 cap (at or above \$215,280 annual or \$103.50 hourly), Methodology must flag for CO review.
6. **Shift coverage upfront in Information to Collect.** Added as Optional Input row so worker asks about 24x7/16x7/12x5 before Step 0.5 is needed.
7. **Methodology depth guidance.** Target 8-12 sections, 2-4 sentences each, readable in 3 minutes. Longer than 14 sections usually means restating Sheet 1-4 data.

## CR-specific patches

1. **SOC 19-1042 Medical Scientist added to Research/Science table.** S1 worker had to improvise between 19-1099 (Life Scientists All Other) and no-match for PhD biomedical. Medical Scientists, Except Epidemiologists is the correct SOC for NIH/pharma PhD biomedical researchers.
2. **Block layout parameterized by fee type.** Sheet 2 block size varies: CPFF = 19 rows, CPAF = 21 rows (adds Base Fee Rate + Award Pool Rate + Assumed Earned %), CPIF = 23 rows (adds Target Fee Rate + Share Ratio Over + Share Ratio Under + Min Fee + Max Fee). Assumption block row ranges: CPFF rows 2-13, CPAF rows 2-15, CPIF rows 2-17.
3. **Asymmetric CPIF share ratio.** Single share\_ratio variable split into contractor\_share\_over and contractor\_share\_under. Real CPIF agreements frequently asymmetric (80/20 over, 50/50 under); exposing both directions separately lets the CO see each leg independently.
4. **CPIF bound-crossing variance.** Baseline  $\pm 10\%$  variance supplemented with  $\pm 25\%$  or wider variance that crosses the min/max fee bounds. Methodology now documents when share ratio stops applying.
5. **CPAF 3-scenario fee view.** Summary shows (a) base only 3% worst, (b) base + 85% pool 8.95% target, (c) base + full pool 10% ceiling. Prevents single-point 85%-earned estimate hiding the range.

## Editorial fixes

1. **Rate Validation status text neutralized.** "Needs explicit justification" replaced with "Position outside  $\pm 25\%$  band; document stacked factors in Methodology" (preserving the  $\pm 25\%$  CR threshold).
2. **Sheet 5 travel skip-or-include contradiction resolved.** Prior copy said "skip if no travel" in one place and "include stub" in another. Now consistent: always include sheet, show "Travel Not Applicable" text when no travel.

3. **Stage A/B skip condition sharpened.** Skip when user provides all four: LCATs with discipline, location with metro, FTE counts, PoP. If any ambiguous or missing, run the gate.
4. **igce\_benchmark promoted to default.** mcp\_\_gsa-calc\_\_igce\_benchmark is now the default tool for Workflow A rate validation; keyword\_search reserved for example-rate or labor-category bucket needs.
5. **NAICS/PSC proactive ask.** Step 0 Stage B parameter question list now explicitly includes NAICS/PSC alongside fee type, metro, contract start.

### **Dropped (too scenario-specific, not worth horizontal ship)**

- BAA cooperative-agreement fee rules (agency-specific, not generalizable)
- OCONUS cost pool adjustments (single OCONUS scenario insufficient data)
- Subcontractor fee-on-fee prohibition edge case (vendor-side rule, not IGCE)
- Termination for Convenience cost estimation (separate activity)
- DCAA forward-pricing rate proposal templates (not IGCE scope)
- Modular budget patterns from NIH R01 (belongs to Grants Budget Builder)
- Limitation on Subcontracting for small business set-asides (pre-award check, not IGCE)
- Award Fee Evaluation Factor weighting (performance monitoring, not IGCE)

### **What has NOT been tested on CR**

- BAA cooperative agreements under FAR 35.016 with non-standard fee structures
- Termination for Convenience cost estimation
- CR-to-FFP conversion modeling (pricing legacy cost-plus scopes as FFP)
- OCONUS CR builds (per diem covers CONUS only)
- 24x7 shift coverage with CR fee math (Step 0.5 untouched since inheritance)
- Custom cost pool rates supplied by CO (skill has rule; no direct test)
- Sonnet 4.6 parity on Wave 2 patches (all runs Opus 4.7)

### **Wave 3 (universal patches derived from CR Wave 2 detailed-prompt round)**

**Wave 3** (Universal patches derived from CR Wave 2 detailed-prompt round): The detailed-prompt scenarios surfaced 31 additional evaluator findings beyond the lazy round. 11 were triaged as univer-

sal-principle patches worth shipping to all 3 IGCE skills; the rest were scenario-specific or architectural (deferred). Patches applied: page\_size=0 deprecation, 24x7 math contradiction resolved, DATEDIF on text cells fixed, day-trip M&IE double-discount corrected (correctness bug shipping 25% low), aged-wage row placement explicit, Sheet 2 hourly vs Sheet 1 annual clarified, BLS flat-tail detection rule, installation crosswalk expanded with 6 DoD/DOE test ranges, SOC 17-2199 fallback documented, same-metro TDY proximity check, stacked factors term enumerated. Status: inherited patches across FFP / LH-TM / CR identically.

## Wave 4 (4 cold tests targeting untested territory, 4 patches shipped, 4 regression tests)

Fourth wave ran 4 cold sub-agent tests on post-Wave-3 skill, then ran 4 regression agents after patches landed. Scenarios were deliberately chosen to exercise items from the Wave 2/3 "not tested" list: CO-supplied DCAA rates, FCCM layer, heavy pass-through ODCs with fee implications, and the CPFF Completion vs Term Form distinction.

### Cold-test scenarios

Scenario	Fee type	Trigger	Finding
GTRI DARPA neuromorphic R&D, \$180K chip samples + \$240K cloud as pass-through	CPFF	CPFF Completion vs Term Form ambiguity + fee-on-ODCs trap	Skill was silent on (d) (1) vs (d)(2); skill implicitly fee'd all ODCs
FEMA Booz Allen CPFF with FPRA (37.8/52.1/11.3/0.22)	CPFF	CO supplies DCAA-audited rates; FCCM non-zero	Skill reverted to 32/80/12 defaults; no FCCM layer
IRS GDIT modernization with \$10.9M 5-year pass-through ODCs (Azure, Snowflake, Splunk, security tools)	CPAF	Fee-on-ODCs trap at scale	Skill would have over-fee'd by ~\$1M across 5 years if fee applied to TEC
NASA KSC Jacobs CPIF with FPRA + FCCM + \$415K/yr pass-through ODCs + asymmetric 75/25 over, 55/45 under	CPIF	All 4 patches in one scenario	Every patch fired correctly

## Patches shipped (Wave 4)

#	Patch	Section affected	Trigger	Severity
1	<b>Fee-Bearing Cost vs Non-Fee Cost split</b> with explicit fee formula applying only to Fee-Bearing	Step 3 cost pool buildup	Every CR build with pass-through ODCs was silently over-fee'd if fee applied to TEC	<b>Critical correctness fix</b>
2	<b>CPFF Completion Form (FAR 16.306(d)(1)) vs Term Form (FAR 16.306(d)(2))</b> prescription with decision heuristic and mandatory citation	Step 3 fee structure, Stage B gate	Skill cited FAR 16.306 without subparagraph; models guessed inconsistently between Completion and Term	Universal structural gap
3	<b>CO-supplied DCAA rates override rule</b> (use FPRA when CO supplies, do not revert to defaults)	Optional Inputs row	Skill had default 32/80/12 but no rule for when CO supplies audited rates	Universal structural gap
4	<b>FCCM (FAR 31.205-10 / CAS 414) as distinct cost pool layer</b> applied to (Subtotal + G&A)	Step 3 buildup, scenario analysis, Optional Inputs	Skill missing this layer entirely; DCAA-audited contractors with CASB Disclosure Statements routinely have non-zero FCCM	Universal structural gap

## Regression validation (4 agents, zero new gaps)

All 4 regression agents ran cold against the post-patch skill on scenarios designed to stress the 4 patches:

- **GTRI DARPA neuromorphic CPFF Term Form, fee-on-ODCs at \$420K pass-through:** Term Form declared in Summary cell B5 with full rationale. Fee =  $7\% \times \$8.88\text{M}$  Fee-Bearing = \$621,922. If fee had been on TEC, would have been \$651,322. Over-fee avoided: \$29,400 across 3 years. Both patches fired.
- **FEMA Booz Allen CPFF with FPRA (37.8/52.1/11.3/0.22):** All 4 FPRA layers applied in order. FCCM visible as distinct layer between G&A and TEC. Methodology Section 4 titled "Cost

Pool Basis: CO-Supplied DCAA Rates per FPRA." Zero reversion to 32/80/12. Both patches fired.

- **IRS GDIT CPAF with \$10.9M 5-year pass-through ODCs:** Fee-Bearing Cost \$42.6M, Non-Fee Cost \$10.9M separated in Summary. Fee at 8.8% target on Fee-Bearing only = \$3.75M. Naive fee on full TEC would have been \$4.71M. Over-fee avoided: \$959,200 across 5 years (lands within 4% of the test spec's ~\$1M benchmark). All 3 CPAF fee scenarios (base/target/ceiling) correctly applied to Fee-Bearing only. Patch fired.
- **NASA KSC Jacobs CPIF with FPRA + FCCM + pass-through ODCs + asymmetric share ratios:** All 4 patches fired simultaneously. FPRA rates (35.2/78.4/9.6/0.31) used. FCCM layered. Fee on Fee-Bearing only, over-fee avoided \$120K across 4 years. Term Form prompt correctly did NOT fire (CPIF, not CPFF). Asymmetric share ratios 75/25 over, 55/45 under held. Bound-crossing variance documented. No regressions.

No new universal structural gaps surfaced. Two minor editorial observations (FCCM-adds-1-row footnote on Sheet 2 block-size table, Stage B gate-skipping criteria could specify "fee form also required") deliberately skipped per universal-only discipline; both are documentation polish that the model handled correctly in practice.

## Bottom line: fee-on-ODCs was the largest silent correctness bug in the CR skill

Across the 4 regression scenarios, the Fee-Bearing vs Non-Fee Cost split prevented a combined \$1.1M of silent over-fee (\$29K + \$959K + \$120K, excluding the FPRA scenario which had no ODCs). Before the patch, any CR IGCE with pass-through ODCs was applying fee to the full Total Estimated Cost, which contradicts standard CR fee practice where fee bears only on contractor execution (labor, burdens, travel) and not on pass-through items. The patch codifies this explicitly with a formula prescription and a CRITICAL call-out.

*Testing record prepared April 2026 by James Jenrette / 1102tools. Four waves documented: Wave 1 inherited, Wave 2 six scenarios (lazy plus detailed prompts), Wave 3 universal patches, Wave 4 targeted untested territory with 4 patches including critical fee-on-ODCs correctness fix. MIT licensed. Source: [github.com/1102tools/federal-contracting-skills](https://github.com/1102tools/federal-contracting-skills).*