Background

K. J. Ross and Associates (KJRA) were engaged to provide User Acceptance Testing (UAT) test management services as well as broader testing process assessment (audit) for and on behalf of QHEST; including:

- Accountability for the integrity of the key outcomes needed from UAT
- Assurance on the Entry criteria being met to commence UAT
- Assurance on the exit criteria being met to commence cut over and go live

Below are the Statuses of the various Activities at various stages:

Test Case Design Review (22/4/09)

OVERALL ASSESSMENT

RED	The test cases do not satisfy requirements for readiness. Areas to be to be immediately addressed are noted below. These areas must be adequately addressed before the project can claim that the test cases identified are ready to be executed.

Audit of draft QHIC System Test and SIT completion report content (24-Apr-09)

Findings

ID	Location	Observation	Commentary
1.	Executive summary, paragraph 2	39 tests are identified as incomplete yet following table show 37 tests incomplete.	Generally throughout document miscounting errors occur due to "Conditional Pass" being sometimes counted as a pass and sometimes as a fail.
2.	Executive summary, paragraph 2	39 incomplete tests are blocked by severity 3 defects which, by definition must have a workaround.	It was explained that at time of review the intention was to resolve the defect rather than apply a work around.
3.	Executive summary, paragraph 2	46 defects are outstanding (per section 3.4.1) blocking 39 tests.	A many to many relationship between defects and incomplete test cases was explained and shown in QC
4.	Section 2.1, Exit Criteria table, Row 1	100% of High Priority tests are claimed to be executed	Prioritisation of test cases was demonstrated however most are rated as High Priority and data in support of the 100% executed claim was unable to be shown
5.	Section 2.1, Exit Criteria table, Row 3	MTP updates are noted as due to be delivered 27 April	This exception to meeting the exit criteria and would need to be accepted by QHEST

6.	Section 2.1, Exit Criteria table, Row 4	Severity 3 defects are proposed to be completely resolved by week 6 of UAT execution on a "best endeavours" basis.	QHEST should plan UAT execution with sufficient contingency to reduce impact of ate delivery
7.	Section 3 Testing Results	Random sampling of QC test sets indicated true reflection of results excepting inconsistent approach to counting of "Conditional Pass" results (refer item 1. above)	IBM to reflect consistent counting of "Conditional Pass" as a passed in final report.
8.	Section 3.2.7 System Test Phase 2	E2E Regression was exploratory testing by SMEs. The 57 tests cover more functionality and risks than other scripted tests.	E2E Regression test is not a repeatable suite. It would be reconfigured for any future use.
9.	Section 3.2.7 System Test Phase 2	Defect 32 was tested using Access based automation tool for WB. Automation expert was unable to demonstrate exact test execution and fail counts for this defect.	QHEST should ensure UAT applies special focus to testing of this functionality.
10.	Section 3.3 System Integration Test	Payslip Advice, Payment Summary Print File interface and QSuper Superannuation interface are noted as to be signed off by tomorrow 24/4.	Payslip Advice was explained as unable to be resolved in current test environment so signoff unlikely for this phase of testing.
11.	Section 3.4.1 Open Defects	Defect counts by severity shown in QC. A singular severity 2 (#1967) is present from these test phases, although others from other phases exist.	An accurate record
12.	Section 3.4.2 Total Defects	Defect counts by severity shown in QC.	An accurate record. Spread of defects amongst severity is indicative of a managed severity assignment process.
13	Section 3.5 Mapping of Requirements	Requirements Traceability Matix V10 was shown. Many requirements were shown to have no test cases listed. Upon examination of QC some test cases were observed without linked requirements.	An urgent reconciliation should be conducted. There is an unacceptable risk that some requirements may have had no testing conducted.
14.	4 Conclusions and recommendations	Defect 1967 was not resolved 23/April as stated.	IBM need to update this section prior to final publication.

		Payslip Advice issue is not able to be proven in existing environments.	QHEST SMEs need to decide on acceptability of the status of the outstanding issues
15.	Appendix A. "Closed with Workaround"	QC records of defects closed in this status matches list provided. QHEST Workaround Register V2.91 show all workarounds and reverse trace to initiating defect	QHEST SMEs need to decide on acceptability of workarounds.

Summary

Draft document indicates a willingness to provide an accurate record. Administrative errors (counting) need to be corrected.

Significantly Issue 13 above, requirements tracing, requires urgent attention and should be considered an unacceptable outcome for test exit.

All findings should be re-examined in light of the final report.

Progress of QHIC User Acceptance Testing (18-May-09)

Current Situation

One quarter of the scheduled UAT execution time has now elapsed.

Progress through test cases has been significantly slower than planned, roughly 30% behind schedule as at 15/5.

- **8 days of execution and currently 3 days behind schedule. Causes are:**
  - **39 newly discovered Sev 2 defects impacting testing execution since start of UAT (17 have been "closed" since May 11th)**
  - **Slow progress of execution due to uncertainty with new processes**

Test case failure rates and defect discovery has been significantly higher than planned, roughly 30% and 9/day respectively. This has included 3 severity 1 defects.

- **278 executed, 200 passed**
- **77 defects total over 8 days of UAT**

The consolidated issue list (sev 2 issues at start of UAT) is falling at a satisfactory rate and appears on track to be resolved within 4-6 weeks but the overall defect number is increasing slowly i.e. slightly more defects are being discovered than closed.

The high number of work-arounds and still resident defects were not factored into the execution planning.

Consequence

It is our opinion that progress will not improve. If the overall defect resolution does not begin to significantly outpace the discovery rate soon it will in fact deteriorate further. At some point all attemptable test cases will have been exhausted. Effectively a stalemate situation.

Recommendations

Acknowledging a strong desire to deliver the system to production as currently scheduled KJRA recommends the following adjustments to the project:

1. Testing as presently sequenced continue. It appears to be effectively identifying the residual defects despite the expectation that most should have been found in earlier testing phases.

2. The defect resolution rate, at this stage of testing, needs to be double the discovery rate. We believe IBM should urgently take whatever capacity and/or efficiency steps necessary to achieve at least this rate.

3. Apply significant resources (already under planning by the Project Manager) to be applied to mitigate error handling and workaround process impacts to the execution rate and finish as scheduled. Note this will also increase the defect discovery rate and without an improved resolution rate would reach a stalemate situation sooner.

4. An additional phase of testing should be scheduled or the regression test phase be extended. A final period of end to end testing needs to occur after code freeze to ensure that an acceptable system is being delivered. The stability assumptions for the present plan have not been met and hence the coverage for this end phase needs to increase.

Audit of QHIC Performance Testing (6-Jul-09)

Findings and Observations

ID	Finding/Observation
1.	**TEST PLAN REVIEW (Conducted in Sept 2008)** **Inspection Item 1** We highlighted the following concerns; 1. Non-functional requirements were not adequately defined. The only reference to a response time target was in regard to Citrix performance. 2. The exit criteria did not define the criteria for determining outstanding defect priority 3. Test Objectives and Pass/fail criteria were not stated in measurable terms 4. There was insufficient test rigour around Payroll Performance Validation and Parallel Payroll Testing given that this was of great concern to the stakeholders 5. Undue focus in the Test Plan on Citrix performance, given that Citrix is a shared environment, and the difficulties that would be introduced when determining sources of poor performance.

ID	Finding/Observation
2.	**PART 1 PERFORMANCE TEST RESULTS (Citrix)** **Inspection Item 2** 1. We agree with the test results and recommendations provided by CorpTech 2. The results are an indication of Citrix performance rather than the end-to-end system, given that only 1/12 of the expected production load was achieved.
3.	**WORKBRAIN LEAVE REQUEST PERFORMANCE** **Inspection Item 3** 1. We agree with the report that the issue relates to the workstation performance and that performance testing is not the appropriate activity to address this issue. 2. Other methods to identify and remediate this issue were suggested by the auditor, eg: Compuware's Application Vantage tool which is a toolset already owned by QHealth. 3. It was also noted that CorpTech's recommendations were not acted on by QHIC at the time it was first highlighted
4.	**PART 2 PERFORMANCE TESTING (SAP & Workbrain)** **Inspection Item 4** 1. Performance Testing was significantly delayed due to their having to share test environment with those conducting PPRT. 2. It would have been better to single stream the activities, letting the higher priority activity go first. 3. If both were of equal priority, then an alternative arrangements should have been considered. 4. Concerns were held concerning the change control process, ie: frequent code drops, which created performance testing rework. 5. This in itself indicated that during the early stages, the application was not sufficiently stable for performance testing. 6. Incidents/Defects discovered during performance testing need to be captured in Quality Center.
5.	**INTERIM PERFORMANCE TEST SUMMARY REVIEW** **Inspection Item 5** 1. We support the performance test findings and recommendations. 2. We agree that CorpTech's concerns about Workbrain performance and scalability should be taken seriously. 3. We agree with the assumptions made by CorpTech regarding response time targets, given the lack of clear measurable non-functional objectives from QHIC. 4. We noticed that our recommendations regarding the Test Plan were not incorporated which meant that CorpTech were reporting against subjective performance objectives. 5. We are concerned that the final version of the summary was submitted without allowing the

ID	Finding/Observation
	auditors opportunity for prior review. 6. The test summary needs to distinguish between QHIC and Whole of Govt objectives. 7. We feel that a full Performance Test Summary Report should be provided to QHIC at this point, given that this concludes the agreed testing according to the revised test schedule. 8. The next round of performance testing (Sept 09?) should be regarded as a separate exercise.

Recommendations

Confidence that this phase of testing is satisfying its objectives should be high.

Our recommendation is that consideration be given to the following:

ID	Description
1.	A Final Performance Test Summary Report should be provided now, ie: at the conclusion of Round 2, incorporating all the performance test results thus far, and modified based on the review comments/suggestions from the auditor.
2.	Allow the performance test auditor to review the Summary Reports from PPV and PPRT testing against their objectives.
3.	Raise Quality Center defects based on the performance issues discovered during testing (if these have not already been raised) to give them appropriate visibility.
4.	Action the recommendations and address the outstanding defects highlighted by Performance Testing so that Round 3 of performance testing can retest the remediation.
5.	Conduct a debrief of performance testing with all interested parties.
6.	Regard Round 3 of QHIC Performance Testing as an additional activity. Review the performance requirements now that actual data is available, assess the items of performance risk, set test objectives to address those risks, conduct test execution (including retests of outstanding performance issues), and provide reporting.
7.	Agree and action ongoing performance test audit activities

Audit of QHIC PPRT (15 Jul 09)

Findings

ID	Finding

1.	**AUDIT PROCESS** These findings are preliminary, the following actions are outstanding: 1. Audit participants should have opportunity to respond/clarify based on preliminary findings. 2. Authorization of scope reduction for retro on retro has not yet been sighted. 3. Further examination of QH Internal Audit findings is desired by this auditor.
2.	**OBJECTIVE** The plan states - *"The focus of the test is on employee payroll results to ensure employees will be paid correctly when the payroll is processed on the new system."* With this intent the planned for accuracy of reconciliation is concerning. Sample size and make up is satisfactory but 10c tolerance per payment type, per employee is a significant variance. Consider: 1. Total payroll approximates $180m per period. Potential under and overpayments per payrun may exceed $40k. 2. Employees will notice payslip variances greater than 1 whole cent and clarification may consume significant resources. The auditor has been advised reconciliation was conducted on the gross not per payment type as planned. This is better than planned but may still necessitate change management. Sighted $134k worth of underpayments and $106k of overpayments greater than $100 in 10% sample from Daily Status Report. This auditor is surprised that Internal audit has not flagged similar concerns.
3.	**TOOLS** Plan states that EPI-USE tool will be employed for verification. An alternate access database was used. Provenance of tool unverified but examination of results contained within indicate likely accurate. All PPRT defects have been closed in IBM QC and moved to QHEST QC instance. Transferred defects are marked as UAT-PPRT. New defects are raised as UAT-OTHER. There is a risk that these defects may erroneously be considered part of UAT scope. UAT is not the appropriate phase to validate large scale reconciliation of payments.
4.	**SCHEDULE** Plan states that 3 cycles of testing would be conducted. It is unclear how many cycles have actually been completed. Exit criteria are being/have been renegotiated. Evidence of approval still to be sighted. Scope for retro testing is being requested for approval on a much smaller subset (10 employees) than planned (100). SSP, based on greater understanding of actual demand, are requesting greater confidence than planned for.
5.	**DATA** The plan states – *"It should be noted that the data conversion process is a critical component to the commencement of Parallel Testing as it creates the fundamental core data used during test execution. Any data conversion errors that are encountered during Parallel will be reported as defects, which must be corrected and re-tested."*

	The production migration approach/tools were not used to convert the configuration data for PPRT. Transactional data was manufactured, and was the only logical option. The volume of defects encountered in early weeks was far larger than expected. IBM sought but did not find a systemic issue with the migration. Individual defects were not originally captured. This is a missed opportunity for improving the production migration. IBM note in the status review *"The data for PPRT was extracted and migrated months ago in the project and had missed the significant improvements in cleansing and migration since then"*
6.	**DEFECT MANAGEMENT** The large number of defects has stressed the defect management process. It was observed that fixes are being applied that undo previous fixes (e.g. QC defect# 3065 appears to revert #949). Once a defect is resolved the record in the access DB is cleared and there is no way to identify where the defect existed. Daily automated queries are being used to monitor regression and are finding instances of previously reconciled records failing after subsequent change. 461 records in the Access DB point to "Lattice Anomaly", the implication is that awards have been interpreted differently over time and therefore these defects cannot be systemically resolved.

Summary

Whilst our initial findings were that confidence that this phase of testing is satisfying its objectives should be low, the subsequent review 22/6/09 has altered that opinion. Whilst elements of the process were sub-optimal we believe the results are satisfactory and we cannot recommend an effective alternative to increase that confidence.

We believe the project should proceed

Our recommendation is that consideration be given to the following:

ID	Description
1.	Audit be ongoing and complete the steps outlined in finding 1 above.
2.	Auditor be re-engaged upon presentation of the test summary by IBM to validate content against plan, earlier findings and current state.
3.	Establish an effective defect triage process to ensure consistent application of fixes e.g. interpretation of an award for a fix for one defect should be the same for any related defect, reducing reappearance of resolved defects. This consistent application of fixes will assist the response to Grandparent Awards.
4.	Plan for Organisational Change Management issues around discrepancies in payslips.

QHIC UAT End to End (4 Aug 2009)

Current Situation

The current and previous attempts at UAT have not achieved the desired outcome of ensuring the new Payroll/HR/Rostering systems of SAP-HR and Workbrain meet the needs of QH. In all instances defect discovery has been significant and the consequent resolution of same has had broad impact to confidence in the system.

Our previous recommendation was to persist with the planned UAT as it was proving effective in discovering defects (over 300 discovered) but acknowledging that the confidence that business needs were met would not be achievable. Further we recommended a final, true, UAT be scheduled. We are pleased these recommendations were accepted and planning for this UAT end to end (E2E) phase is now well advanced.

To succeed as a true UAT, building confidence that the system is ready for production use, the E2E test needs to commence with a system of representative quality to that intended for production use. That is why current entry criteria state that all severity 1 and 2 defects must be fixed and retested prior to test commencement. However, current defect fix and retest trends indicate that this will not be achieved by the scheduled commencement date and impact analysis is underway.

To ensure a speedy UAT E2E, in support of an imminent go-live, additional shifts of testers are being considered.

The Concerns/Risks

1. Defect Discovery was still trending upwards at conclusion of testing and high severity (1 and 2) had not tapered off. In our experience this indicates significant volume of defects remain undiscovered.

2. UAT E2E planning needs to match the actual entry and exit criteria. If UAT E2E must commence with a defect burden that does not match the planning then the progress of testing will be slower with the following possible consequences: scope of testing will be reduced; schedule overrun; budget overrun; inappropriate/ill-considered defect resolution.

Recommendations

In essence there is only one recommended course of action. That being to commence UAT E2E with a defect burden equivalent to that QH would accept for the production/go-live commencement. But how to most effectively achieve that?

We endorse the current plan for a defect burden for UAT E2E and for go-live of zero severity 1 and 2 defects. If that is not achievable the confession must therefore be that some number of employees can be acceptably paid incorrectly.

Sub-recommendations therefore are:

1. The impact analysis should directed to only allow through those defects which, unresolved, will result in "acceptable" incorrect payment (consider numbers of impacted employees, under vs overpayment, dollar variance tolerance by individual and total payrun). By whatever process this is to be worked around in production should be tested in this UAT E2E. Whatever criteria are used and the per defect ruling against that criteria to be captured in the official record with appropriate and clear approvals. We fear subsequent interrogation of this decision will not be viewed favorably.

2. Consider moving to multiple shifts for the retesting effort in advance of the UAT E2E in order to meet the current planned defect burden. If necessary, delaying the commencement of UAT E2E.

3. Consider allowing either more time or more budget for exploratory (un-scripted) testing during UAT E2E. This could be done at the conclusion of running the planned UAT E2E scripts (more schedule). Alternatively, adding additional resources to execute the scripted component of testing, freeing up the experienced testers for exploratory testing in parallel (more budget).

4. Consider what leverage may be achieved by payment of success bonuses rather than variations. This could be applied if defect resolution rates fall behind or for early resolution of critical defects. The prospect of these payments may enable resolver groups to acquire additional resources and/or the prospect of losing the payment if not delivered on time joins them to your risk of late completion.

5. Obviously, consideration could also be given to delaying go-live and all precedent activities to match current resourcing delivery trends.

QHIC UAT End to End (29 Sep 2009)

Current Situation

The current UAT is scheduled to finish 5th October, however given the current defect trends and test case execution velocity, will not meet the defined exit criteria of ensuring the new Payroll/HR/Rostering systems of SAP-HR and Workbrain meet the needs of QH.

Defect discovery has consistently not shown a trend that would suggest that the system is in a stable state which is a key entry point to the UAT phase.

The current UAT phase was initially scoped at 8 weeks effort but was compressed to a 4 weeks effort to meet a business driven go-live date. There were a number of pre-requisites that needed to be considered before agreeing the new 4 weeks duration:

- A de-scoping exercise of the Finance area test scripts made it a possibility to meet the more aggressive time schedule.
    - Further de-scoping at this stage is most likely not a viable alternative.
- The ability to execute UAT in 4 weeks was completely dependent on the UAT running smoothly with minimal defects and issues with pay runs
    - The escalation criteria was met for Severity1s (2) in first 2 days and for Severity 2s (40) in the first week of execution
- In finance, the test scripts are likely to be only 95% completed in time remaining
    - Key Corptech resources have been lost to the UAT execution in the last week of testing through non-extension of their contracts – they were focused on the ALCS functionality

Given the current situation and statements of fact, it is obvious the UAT execution (including defect retesting) will not be able to be completed in accordance with the stated exit criteria in 4 week timeframe

The Concerns/Risks

1. Defect Discovery is still not trending downwards towards the conclusion of testing and severity 2s have not tapered off. Given the already reduced scope and the fact that by design the UAT effort doesn't generally provide full coverage of system functionality, one could make a fair assumption that numerous defects still remain.

2. Pay run stability is also a concern in that only 1 of 9 executed pay cycles have been successful. There are only 2 pay cycles that could be run in the remaining time allotted to this UAT.

3. Defect reintegration continues to be an ongoing concern. The risk involved in promoting all Severity 2 Priority 1 defects into UAT in the last week introduces and unacceptably high risk.

   Exit criteria in part states that:

   - We must test all scripts
   - We must retest all defect fixes

   There are currently 81 Severity 2 Priority 1 defects currently in Quality Control. 87 fixes are already in the UAT environment which must be retested before exiting. In total there are 168 defects requiring retest prior to phase end on the 4th October.

   There are approximately 3 test scripts required per defect which leaves 504 test cases to be executed for retesting alone. In addition there are the 300 test scripts in the original plan still to be executed.

Recommendations

Given that we do not believe that the UAT E2E will achieve the defined exit criteria for the phase we would suggest that it would be prudent to plan for an additional UAT phase as outlined below:

- There are 100 defects in Red Line (Open with IBM) plus 87 fixes in the UAT environment which means 187 defecst required to be retested

- Based on current trend we would expect to find 40 more defects by the end of current UAT which brings our total to 230 defects to be tested

- At 3 test scripts per defect on average we would require 700 scripts for retesting

- Current throughput of scripts is running at 100/day which would equate to 10 days effort (including contingency)

- There would be a final requirement for 15 days regression testing

Therefore it would require a total effort of 5 weeks for UAT to have met the exit criteria and therefore have the confidence to go-live.

The key to achieving a successful UAT E2E testing phase that meets the exit criteria is totally reliant on a system that is stable prior to entering each stage e.g stable before starting restest, and stable before starting regression.

Definition of stable system: no known open severity 2 defects