Defining and Implementing Metrics for Project Risk Reduction

Tom Kendrick, Program Manager, Hewlett-Packard

Abstract:

"You can't manage what you can't measure"—David Packard, HP Founder

Effective project risk management, like project management overall, depends on measurement. This paper explores uses for all three types of project metrics:

- **Predictive** metrics: forward-looking, based on expectations
- **Diagnostic** metrics: drawn from current project status, throughout the work
- **Retrospective** metrics: backward-looking, derived from results

Following a survey that includes representative metrics in each of these categories, you will find tips for defining a system of useful project measures to improve your risk management. Building your basic set of metrics need not be difficult, and it can make the difference between project success and failure.

Metrics and risk

Useful metrics always have three properties. They:
- Support larger objectives
- Influence behavior
- Assist good decision-making

Metrics related to discovery and minimization of risk directly relate to the project leader's goal of a successful project. How people choose to work greatly affects the risks that a project faces; measurements that reinforce desired behaviors will have a significant effect on staff motivation and project progress. Metrics that improve the quality of project decision-making also contribute to lower overall risk.

Much project risk arises as a result of overconstrained projects. Risk is present whenever project objectives are unrealistic due to scope commitments that exceed the capacity of the project team (or deadlines are too short, or resources are insufficient, which are really the same thing). Undertaking too many projects in an organization also creates risk. One reason that this kind of exposure is so common is that the "I want more" sponsoring and stakeholder faction in most organizations has more (generally much more) authority and position power than the project leader/project staffing community. In setting project objectives and making portfolio choices, the top-down view always trumps the bottom-up.

While there can never be a truly level playing field in this environment, bringing objective data based on project metrics into decisions, discussions, and negotiations makes a huge difference. When a sponsor asks, "Why can't you do more?", project metrics provide a basis for answers that are difficult to ignore. Each of the three types of metrics has a role in project risk management. Predictive metrics use planning information early to identify risks. Diagnostic metrics interpret current status, and serve as timely triggers for risk response and necessary changes. Retrospective metrics assess how well things worked and provide insight into unanticipated risk and recurring problems, enabling better future decisions.

Predictive project metrics

Predictive project metrics serve as a distant early warning system for project risk. These metrics use forecast information, normally assessed in the early stages of work, to make unrealistic assumptions, significant potential problems, and other project risk sources visible. Because they are primarily based on speculative rather than empirical data, predictive metrics are generally the least precise of the three types. Predictive project measures support risk management in a number of ways:

- Determining project scale
- Identifying the need for risk mitigation and other project plan revisions
- Determining situations that require contingency planning
- Justifying schedule and budget reserves
- Supporting project portfolio decisions and validating relative project priorities

Diagnostic project metrics

It's said that a frog dropped into boiling water will hop out promptly, but a frog set in cool water that is gradually heated will sit there until it is too late. Project leaders too often find themselves in hot water for similar reasons; a reasonable-sounding project gradually becomes impossible due to incremental changes in scope, resources, and timing—with no one realizing the shift. Diagnostic metrics are designed to provide real-time information about a system, and they serve as a thermometer for assessing just how hot the water is becoming. Based on project status information, diagnostic project metrics assess the current state of an ongoing project. Risk-related uses include:

- Triggering risk responses and other adaptive actions
- Assessing the impact of project changes
- Providing early warning for potential future problems
- Determining the need to update contingency plans or develop new ones
- Deciding when to modify (or cancel) projects

Retrospective project metrics

Retrospective metrics determine how well a process worked after it completes. They are the project environment's rear-view mirror. Backward-looking project metrics assess the overall effectiveness and efficiency of project processes when a project has finished (or has been canceled). Use retrospective project metrics to:

- Track trends
- Identify recurring sources of risk
- Set standards for reserves (schedule and/or budget)
- Determine empirical expectations for "unknown" project risk
- Decide when to improve or replace current project processes

- Validate the accuracy of predictive metrics and adjust the processes (such as estimating) used to develop them

Defining metrics for risk management

Before deciding what to measure for your project, you should consider behavior changes necessary to improve your management of risk. If past projects have run into difficulty due to excessive scope change, minimizing unnecessary changes will help. To avoid schedule problems resulting from required work that was overlooked in planning, more thorough analysis will have benefits. For resource risk arising from cost overruns, seeking better data for early estimates will minimize surprises.

Begin identifying your key metrics by listing any behavior changes that will affect project risk. Brainstorm with appropriate project stakeholders, developing candidate measurements with potential to encourage the behaviors you desire. If out-of-control project change is an issue, consider tracking the frequency of scope additions and removals as a diagnostic metric during projects, or as a retrospective metric to compare related projects at completion. For unanticipated project work, you might measure the number of activities added to the project after setting the baseline (a retrospective metric, evaluated at project end). For estimation accuracy, a possible metric might be "Cumulative difference between estimated and actual costs of completed project work," a diagnostic metric that is part of Earned Value analysis. Develop a list of behavioral goals that relate to risk management, and use this as a guide in choosing metrics that show how well your goals are met.

A project is a complex system, so implementing too few risk-related metrics will usually not be very effective. Selecting too many metrics is also problematic, as the overhead and cost of collection will be high and important information may be lost in the jumble. Strive to define a sufficient set of project metrics that will provide you with a foundation for effective risk management.

<u>Examples of predictive project metrics:</u>

- Project size/scale risk
    - Project duration (elapsed calendar time)
    - Total effort (sum of all activity effort estimates)
    - Total cost (budget at completion)
    - Size-based deliverable analysis (component counts, number of major deliverables, lines of noncommented code, blocks on system diagrams)
    - Staff size (full-time equivalent and/or total individuals)
    - Number of planned activities
    - Total length (sum of all activity durations if executed sequentially)
    - Logical length (maximum number of activities on a single network path)
    - Logical width (maximum number of parallel paths)
- Scope risk
    - Project complexity (interfaces, algorithmic assessments, technical or architecture analysis)

    - Volume of anticipated changes
- Schedule risk
    - Activity duration estimates compared with worst-case duration estimates
    - Number of critical (or near-critical) paths in project network
    - Logical project complexity (the ratio of activity dependencies to activities)
    - Maximum number of predecessors for any milestone
    - Total number of external predecessor dependencies
    - Project independence (ratio of internal dependencies to all dependencies)
    - Total float (sum of total project activity float)
    - Project density (ratio of total length to total length plus total float)
- Resource risk
    - Activity cost (or effort) estimates compared with worst-case resource estimates
    - Number of unidentified activity owners
    - Number of staff not yet assigned or hired
    - Number of activity owners with no identified backup
    - Expected staff turnover
    - Number of geographically separate sites
- Financial risk—Expected return on investment (ROI)
    - Payback analysis
    - Net present value
    - Internal rate of return
- General risk
    - Number of identified risks
    - Quantitative (and qualitative) risk assessments (severity analysis)
    - Adjusted total effort (project appraisal: comparing baseline plan with completed similar projects, adjusting for significant differences)
    - Survey-based risk assessment (summarized risk data collected from project staff, using selected assessment questions)
    - Aggregated overall schedule risk (or aggregated worst-case duration estimates)
    - Aggregated resource risk (or aggregated worst-case cost estimates)
    - Dilbert Correlation Factor (Collect 30 recent Dilbert cartoons and circulate to staff. Have people mark each one that reminds them of your organization. If the team average is under 10: Low organization risk. 10-20: Time for some process improvement. Over 20: Hire a cartoonist and make your fortune….)

<u>Examples of diagnostic project metrics:</u>

- Scope risk
    - Results of tests, inspections, reviews, and walkthroughs
    - Number and magnitude of approved scope changes
- Schedule risk
    - Key milestones missed
    - Critical path activity slippage
    - Cumulative project slippage

    - Number of added activities
    - Early activity completions
    - Activity closure index: the ratio of activities closed in the project so far to the number expected
- Resource risk
    - Excess consumption of effort or funds
    - Amount of unplanned overtime
    - All earned value management (EVM) metrics (EV, AC, PV, CV, SV, CPI, SVI, and the rest of the alphabet soup)
- Overall risk
    - Risks added after project baseline setting
    - Issues opened and closed
    - Communication metrics, such as volumes of email and voicemail
    - The number of unanticipated project meetings
    - Impact on other projects.
    - Risk closure index (ratio of risks closed in a project divided by an expected number based on history)

<u>Examples of retrospective project metrics:</u>

- Scope risk
    - Number of accepted changes
    - Number of defects (number, severity)
    - Actual “size” of project deliverable analysis (components, lines of noncommented code, system interfaces)
    - Performance of deliverables compared to project objectives
- Schedule risk
    - Actual durations compared to planned schedule
    - Number of new unplanned activities
    - Number of missed major milestones
    - Assessment of duration estimation accuracy
- Resource risk
    - Actual budget compared to planned budget
    - Total project effort
    - Cumulative overtime
    - Assessment of effort estimation accuracy
    - Life-cycle phase effort percentages
    - Added staff
    - Staff turnover
    - Performance to standard estimates for standardized project activities
    - Variances in travel, communications, equipment, outsourcing, or other expense subcategories
- Overall risk
    - Late project defect correction effort as a percentage of total effort
    - Number of project risks encountered
    - Project issues tracked and closed
    - Actual measured ROI

Implementing metrics for risk management

No one could ever justify collecting project data for every conceivable metric that relates to project risk. To select a useful subset from all the possible measures, work with project stakeholders to determine the root causes of your most significant recent problems. Then, identify measures that you believe align most closely with your listed problem sources. Different projects will face different challenges, so no single combination of measurements can fit every situation.

Prioritize any metrics you are considering, using criteria such as criticality, contribution to potential process improvement, linkage to desired behaviors, or availability of data. Collect only metrics that will make a meaningful difference; do not collect data just because you can. For every metric, define how it will be used, and get commitment that it will be acted upon in this way.

An effective set of metrics provides tension—improvement of one measure may diminish another one. Opposing a metric measuring speed of execution with another measuring defects or quality will result in more appropriate behavior than either measurement by itself. Work to minimize "gaming" of the metrics by eliminating factors that might improve the measurement without achieving any desired results. It is possible to subvert almost any metric, so define metrics in terms that minimize differing interpretations and loopholes.

Whatever combination of measures you choose, work to ensure that they are:

- Objective (different people collecting the data will deliver consistent results)
- Accessible (measures are easy to collect)
- Clear (all involved understand the measurement process)
- Important (all metrics collected will be used—not just collected because you can)
- Accepted (people affected by the measurements approve of and support how the measures will be used)

For each diagnostic metric, define a frequency for collection that supports your objectives—frequent enough to support the results you desire, but not so often that it represents high overhead. For metrics that are not simple counts or ratios, define specific units of measure (days, Euros, effort-months) for all data collected to ensure consistent precision and comparability.

For project metrics measuring factors that are under the control of the project team, use input and computational definitions that are unambiguous and not subject to change. Avoid metrics based on subjective interpretations.

Finally, work to ensure that any metrics collected are used primarily for process monitoring and improvement, not as a basis for punishment. Metrics are powerful tools for identifying opportunities for beneficial change and determining trends, but the quality of the data that people provide will be less useful if they know that the data will also be used to evaluate their performance. Once metrics are identified with processes used to rank and cancel projects, the reliability of future data deteriorates substantially. Use metrics for process control and improvement, not to generate criticism of the project team. If any personal information is involved, ensure that the measurements are kept confidential.

Define and document each project metric clearly to minimize differing interpretations in a metric data sheet. Include information such as: the name of the metric, the intended objective, required data, measurement units, frequency, collection method, any formulas used, the target acceptable range, who will make the measurement, and how it will be achieved. An example datasheet:

Metric Data Sheet: Activity Closure Rate	
Objective:	Provide project progress information
Type:	Diagnostic
Normal range:	.95 to 1.1 (higher is better)
Tension:	Output quality, deliverable cost
Calculation:	(# of activities closed) / (Total #) / (% of project timeline consumed)
Data:	Activities complete, current date
Reported by:	Activity owner
Frequency:	Weekly
Tools used:	MS Project (collection and storage)
Potential barrier:	Performing easy, short activities first

Once a set of metrics is defined, the next step is to define an acceptable or desirable normal range. For well-established metrics, baselines are probably already documented. For new measures, or for metrics used in a new application, you need to establish the initial data range. While you can begin with an educated guess as a provisional baseline for a new metric, you should use the first several cycles of data collected to confirm it. Use this initial data only to validate or to correct the baseline. Validate a baseline for the measurements before you make any decisions or changes. Until you have a validated baseline, measurements will be hard to interpret, and you will not be able to determine the effects of process modifications that you make.

Before you start to use any metric, discuss it with everyone who will be affected by it. For each metric, work to get consensus from all members of the project team on the definition, the planned collection and use of the data, and the meaning of the results. Get commitment from everyone who will collect or supply data in advance, and seek agreement not to "game" the metrics.

After setting a measurement baseline, collect project data as planned, and use the information to guide your project decisions. Throughout the process, make the measurements visible. Report the status of measured factors as planned, to all project stakeholders who need the measurements or are affected by them. Be prompt in evaluating and reporting the data to ensure timely feedback and early detection of significant variances.

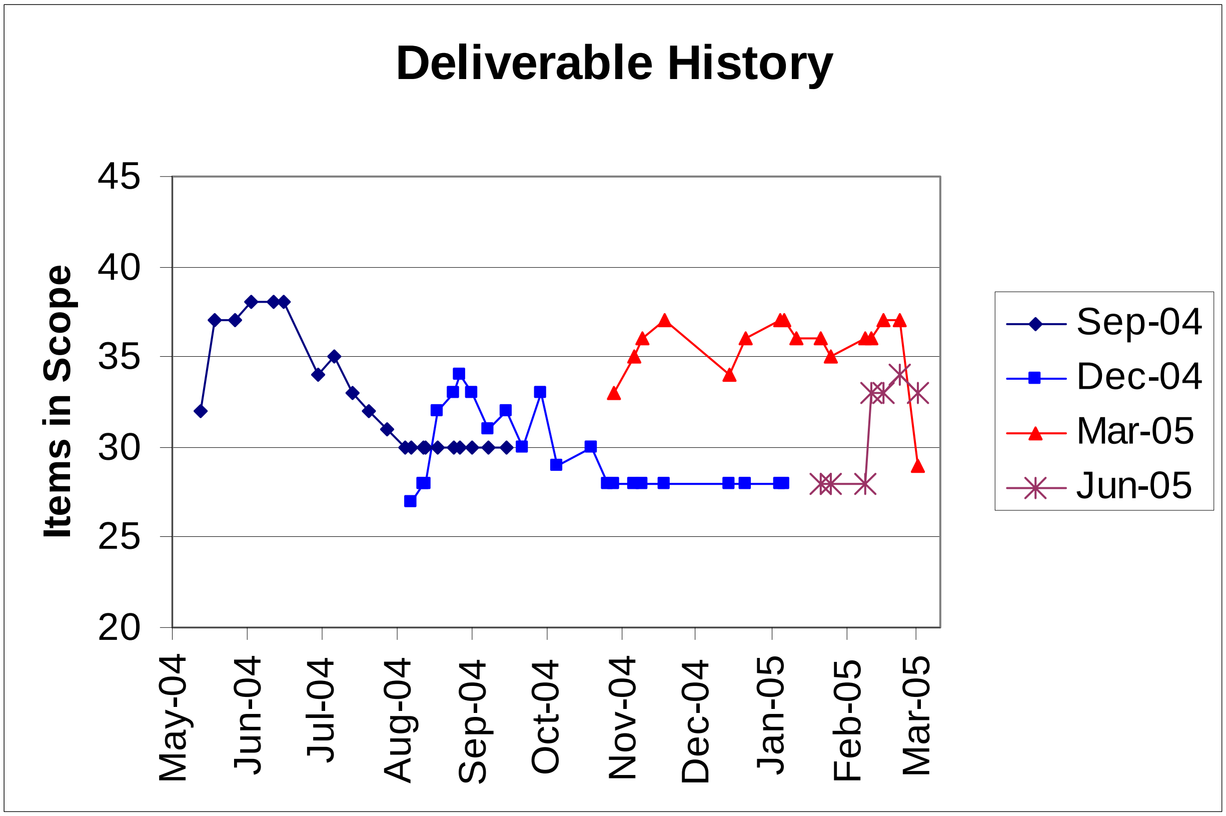
Particularly with new metrics, remain skeptical. Review the data, and confront any suspected “gaming” of reported measurements. Periodically reevaluate all metrics, especially after significant organizational or process changes. Following changes, review the baseline and acceptable range for each metric. Validate any necessary adjustments with new baseline measurements before considering additional system changes.

Example metrics

The following metrics are used for a global Information Technology program with an extended team of about 150 people, located in about a dozen countries world-wide. The program is responsible for global system deployment through about a dozen quarterly waves implementing the system in an ever-expanding list of countries over a period of several years.

<u>Predictive metrics</u>

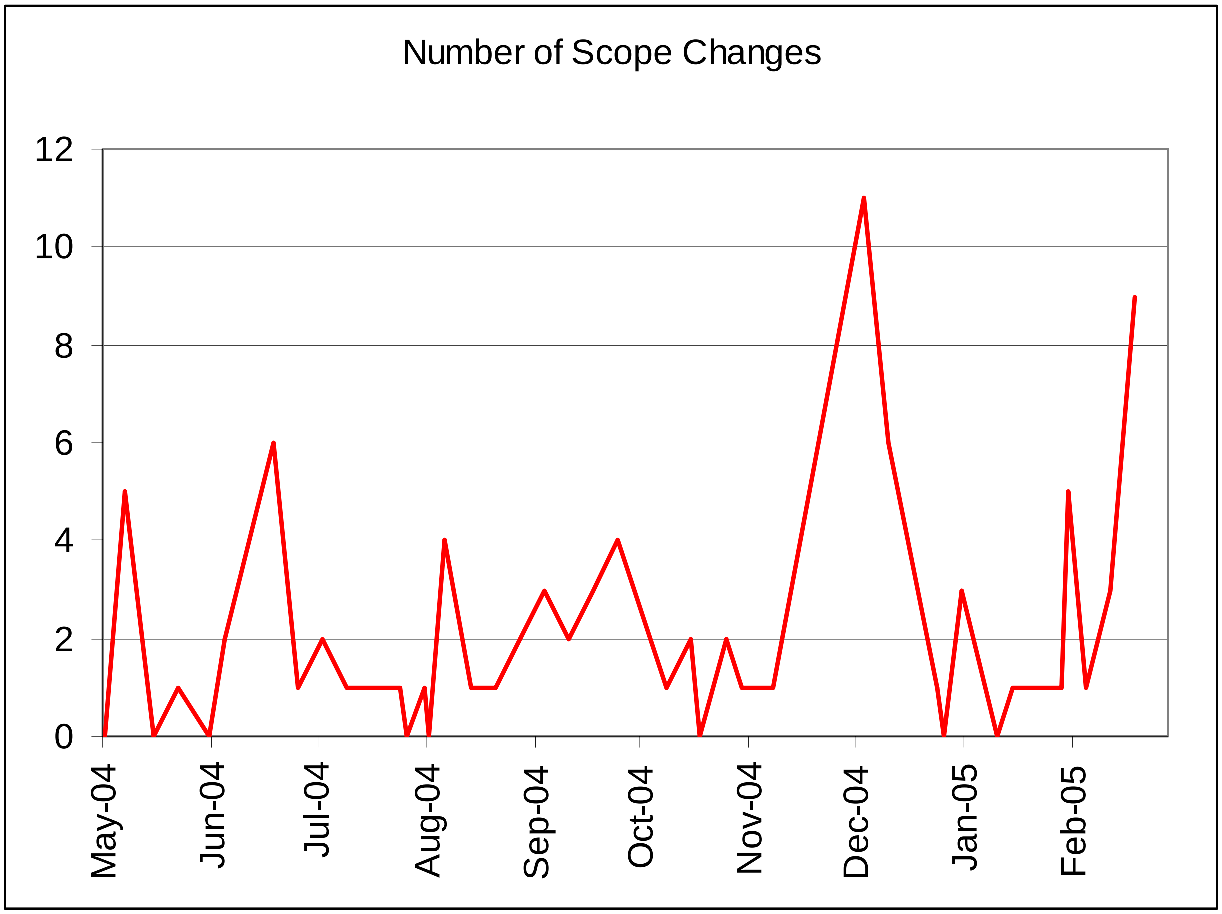
One of the most reliable predictors of project problems is overall scale; the bigger a project is, the more failure modes there are and the easier it becomes to overlook potential trouble. Each deployment wave for this program is staffed and funded similarly, and each of these sub-projects has essentially the same duration from start to finish. Most measurements of scale based on planning are relatively consistent, but one that can vary is the number of specific, independent deliverables included in each wave. The number for the last several cycles has begun and ended at roughly 30, with mid-project excursions approaching 40. A graphic summary covering the past year, connecting the scoping for each wave and identified by its endpoint, follows:



As a predictive metric, the initial and running totals illuminate potential scope risk. Viewed retrospectively, these measures permit more realistic future scoping decisions.

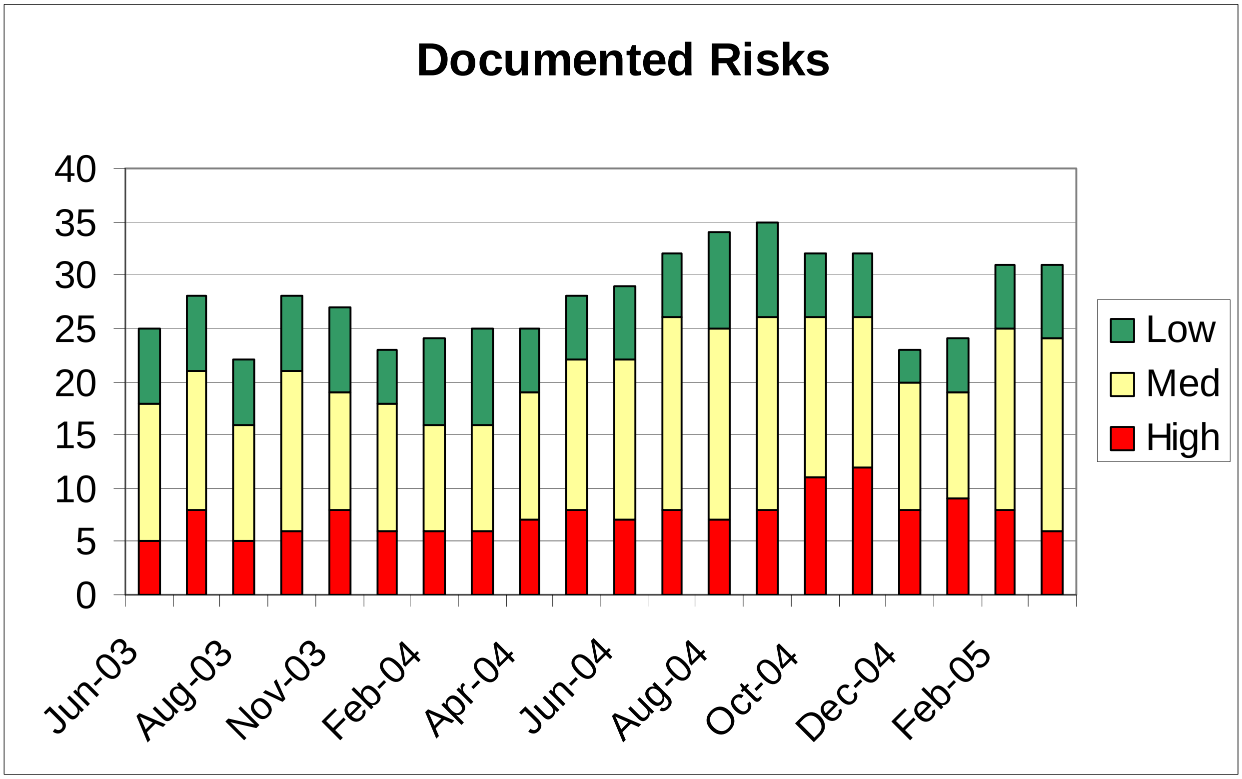
<u>Diagnostic Metrics</u>

In addition to the usual array of status indicators for progress, resource use, and other factors, we also track the number of changes made to scoping. The following diagnostic data corresponds to the previous chart (quarterly scope freeze decisions, which would insert quarterly spikes of about thirty into this chart, are intentionally omitted from this data):

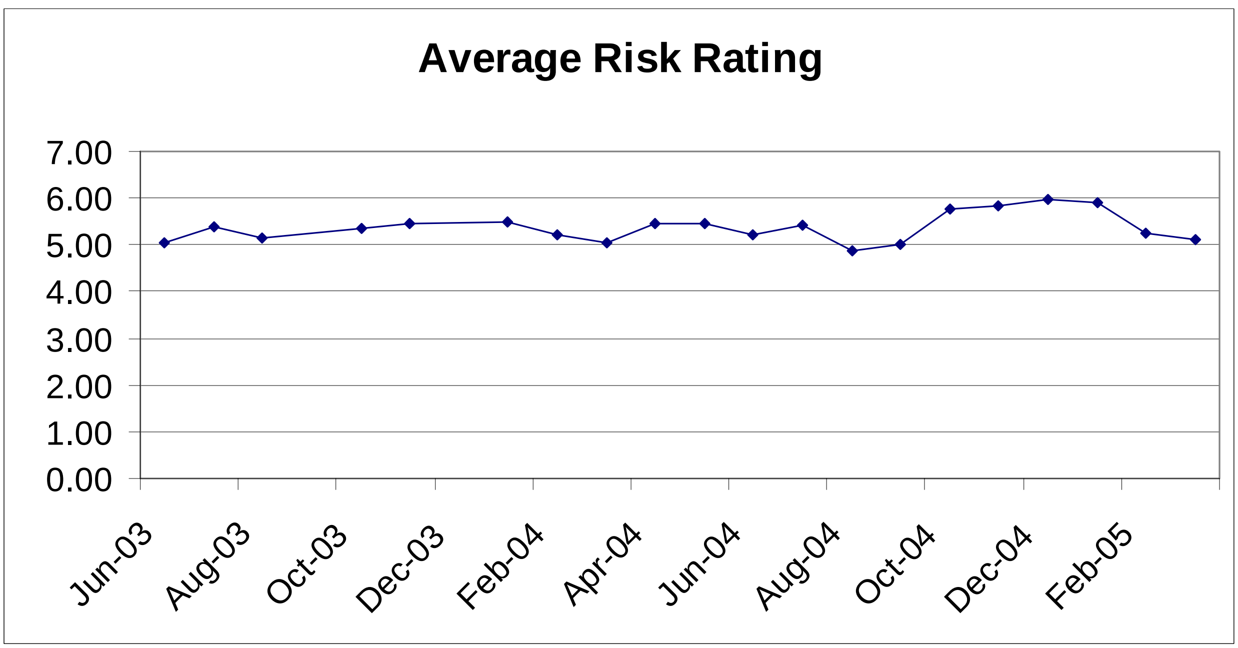


Ideally, the number of changes on average over time should remain constant, or drop. An accelerating rate of scope change represents additional risk to be managed.

We also maintain a running analysis of risks, based on monthly reviews at the program level. The graph that follows shows the number of known program risks and their severity over the past two years:



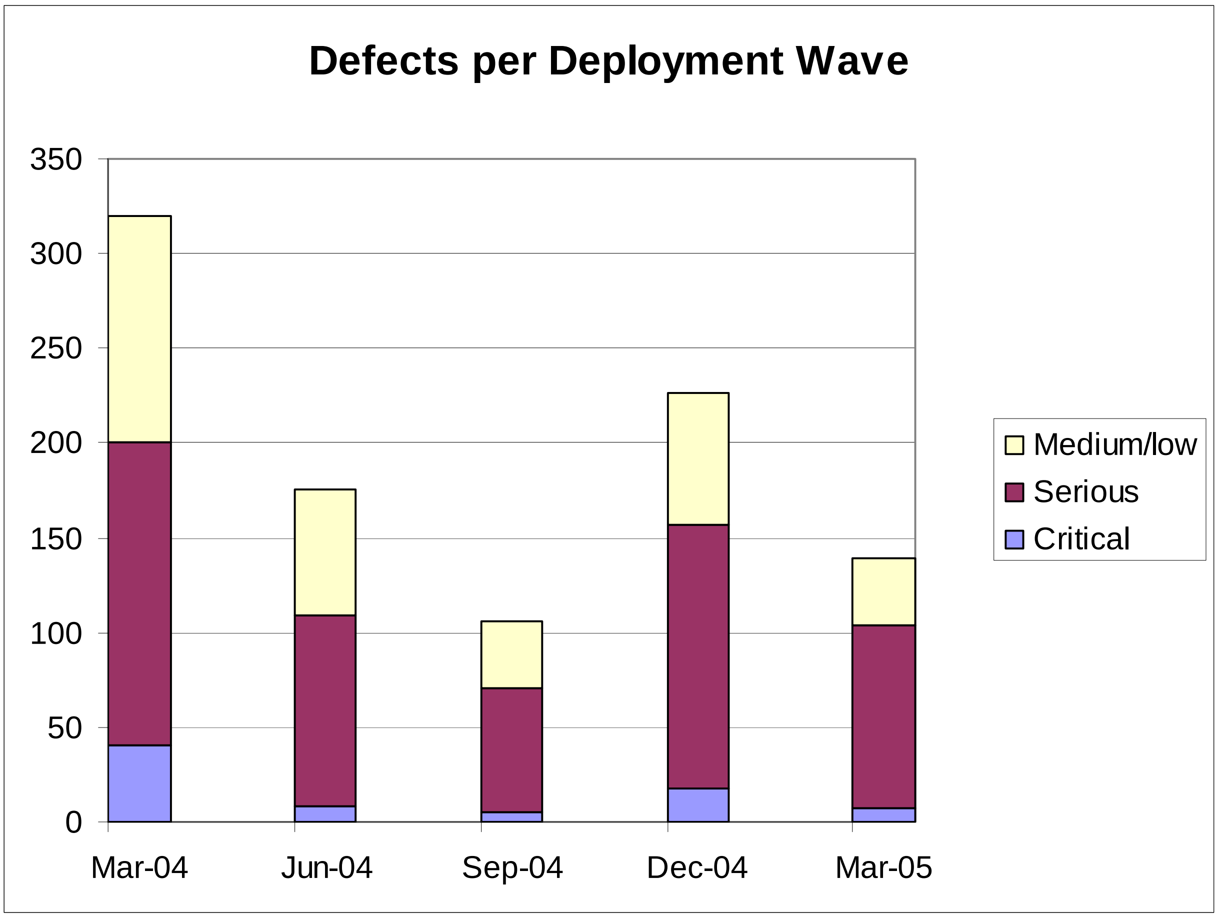
This graph summarizes the typical severity (where “High” is 9):



Risk assessment over time for a lengthy program might be expected to drop, both in number and severity. The profile here does show that things are under control, with a consistent average rating of Medium.

Retrospective metrics

Following each completed deployment wave, a number of overall measures are evaluated, such as scope delivered (summarized in the earlier chart), performance to schedule, and other typical project factors. One additional measure of overall quality and scope management is the defect count for each wave subproject, which is summarized in this chart:



For a mature program well over a year into its run, this kind of measure should be stable or dropping. The stability displayed is consistent with the overall scope of this work, and does display solid evidence of overall process improvement and control.

Conclusion

Managing risk on a project includes shining a bright light into the often murky corners where problems can hide. Thorough planning and analysis of predictive metrics, with particular focus on significant differences compared to prior work, will reveal project risks when they can be most easily managed through plan modifications, changes to the project, or other means. Diagnostic metrics serve as triggers for risk response and as early warning signals of imminent trouble. Retrospective metrics provide a longer view and facilitate discovery of adverse trends, allowing you to make adjustments to current and future work to make it more failure-proof. Metrics also provide the objective data that project leaders need for principled negotiation with project sponsors in order to avoid project risk. Predictive metrics, especially those validated using retrospective data, are a project manager's best defense against unrealistic management requests.

Overall, using project metrics raises the visibility of what is going on, with increased attention to risk management. As HP founder Bill Hewlett was fond of saying, "What gets measured gets done."

Some references on Metrics and Risk Management

DeMarco, Tom. *Controlling Software Projects*. Prentice Hall PTR, 1986. ISBN: 0131717111

Grady, Robert B. *Successful Software Process Improvement*. Prentice Hall PTR, 1997. ISBN: 0136266231

Kendrick, Tom. *Identifying and Managing Project Risk: Essential Tools for Failure-Proofing Your Project*. AMACOM, 2003. ISBN: 0814407617

Kendrick, Tom. *Project Management Tool Kit, The: 100 Tips and Techniques for Getting the Job Done Right*. AMACOM, 2004. ISBN**:** 0814408109

Patterson, Marvin. *Leading Product Innovation : Accelerating Growth in a Product-Based Business*. Wiley, 1999. ISBN: 0471345172