Formal Peer Review

Keith Roseberry, Collins Aerospace
Mary Ann Sheppard, NAVAIR
Robert Wallis, Leidos
Ye Yang, Stevens Institute of Technology

CSER 2019
conference on systems engineering research

Formal Peer Review

Industrial Case Study

Keith Roseberry, Collins Aerospace, Raleigh NC

Mary Ann Sheppard, NAVAIR, Lakehurst, NJ

Robert Wallis, Leidos, Gaithersburg, MD

Ye Yang, Stevens Institute of Technology, Hoboken, NJ

Background & Objectives

- Background
  - Most studies focused on code reviews and open source community
  - Lack of data from diverse disciplines, large-scale industrial environment
  - Perception is that all data must be peer reviewed
- Objectives
  - Establish metrics for determining FPR effectiveness
  - Determine relative effectiveness of FPR practices
  - Evaluate feasibility of FPR practices based on effectiveness measures
    - Is the high cost of FPRs justified for all scenarios?

Introduction

- Formal Peer Review (FPR)
  - Formal process of reviewing output of development/verification activity
  - Benefits are well-accepted
  - Practices vary widely and are inconsistent
  - Laborious and costly
    - Meeting-based FPRs costlier with no demonstrated higher defect detection
    - Hinders wide adoption of FPR practice
- Main Results
  - Little correlation between labor and defects
  - Early FPRs much more effective
  - Some boundary conditions encountered during analysis

Study Process

1. Goal-Question-Metric (GQM) to determine measures to determine FPR effectiveness
2. Define effectiveness model
3. Analyze 100 FPRs from safety-critical system development company
4. Further analysis of 197 FPRs with subset of measures
5. Identify conclusions and further studies

Goal-Question-Metric (GQM)

Goal: Do the results of a FPR justify the cost spent performing the FPR?

Questions

Q1: What is the cost of a FPR?

Q2: What is the effectiveness of a FPR?

Q3: Is there any relationship between FPR effectiveness and review type?

Q4: Is there any relationship between FPR effectiveness and review size?

Metrics

- **M1**: Review Type
- **M2**: Defect Count
- **M3**: Labor Hours
- **M4**: Review Size

- M1 needs to be quantified

GQM M1 (Review Type) Factor

- Assigns a factor to each review type
  - Weighs the cost of fixing a defect when allowed to propagate into the product
  - The code is the product, therefore cost = 1.0 for code defects

Phase	Review Type	Cost Factor	M1 Factor
Requirements	Requirements	1X	10
Design	Design	5X – 7X	2
Build	Code	10X – 26X	1
Test	Test	50X – 177X	0.237
N/A	Document	N/A	0.2
N/A	Configuration Accounting	N/A	0.1

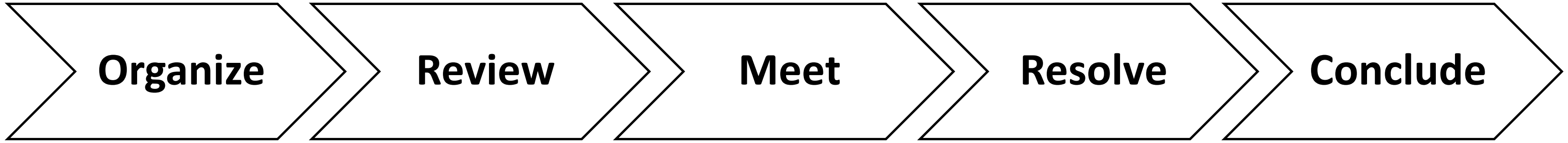
FPR Effectiveness Model

- Simple model to quantify effectiveness
- Future studies may refine the model

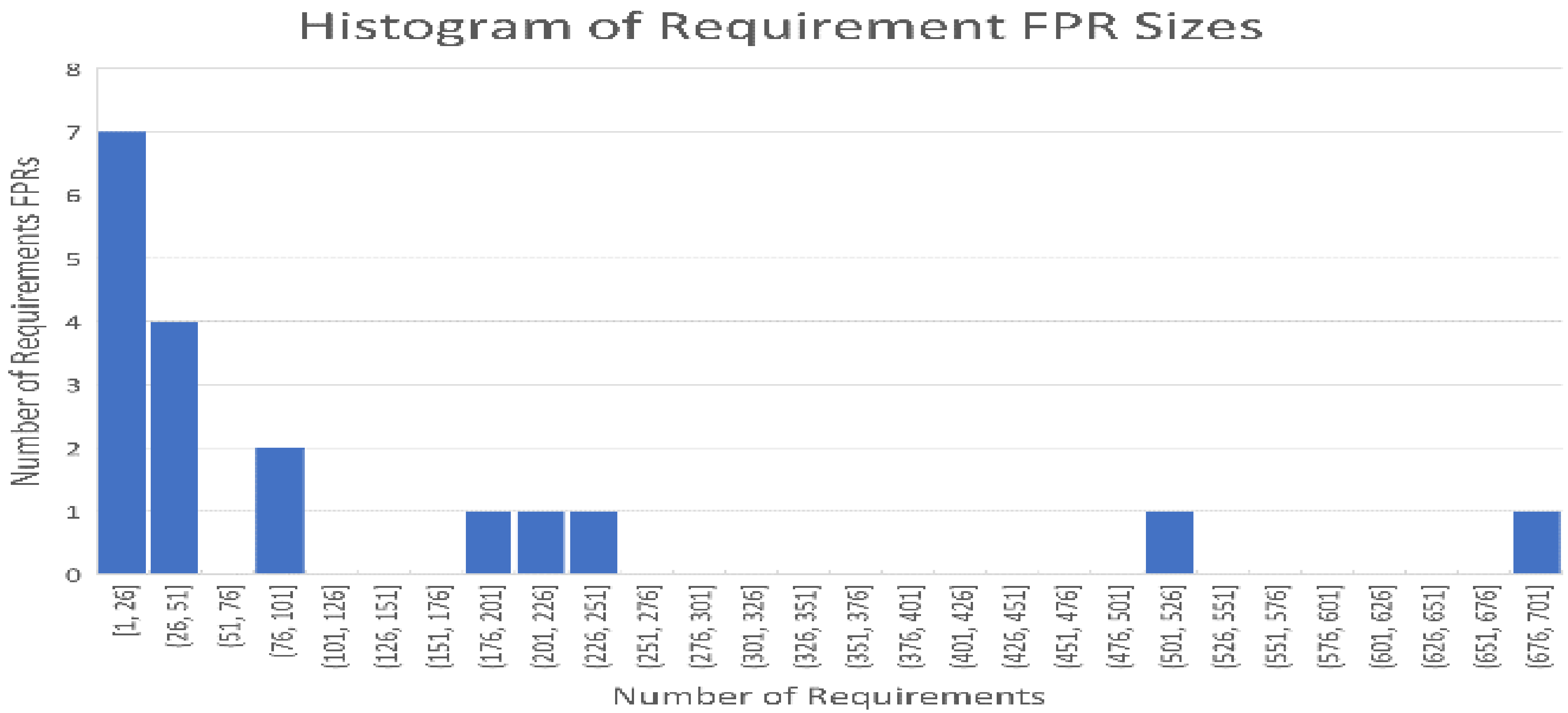
$$\text{FPR\_Eff} = \left(\frac{Defects}{Labor\ Hours}\right) \times \text{Review Type Factor}$$

Company Background & FPR Process

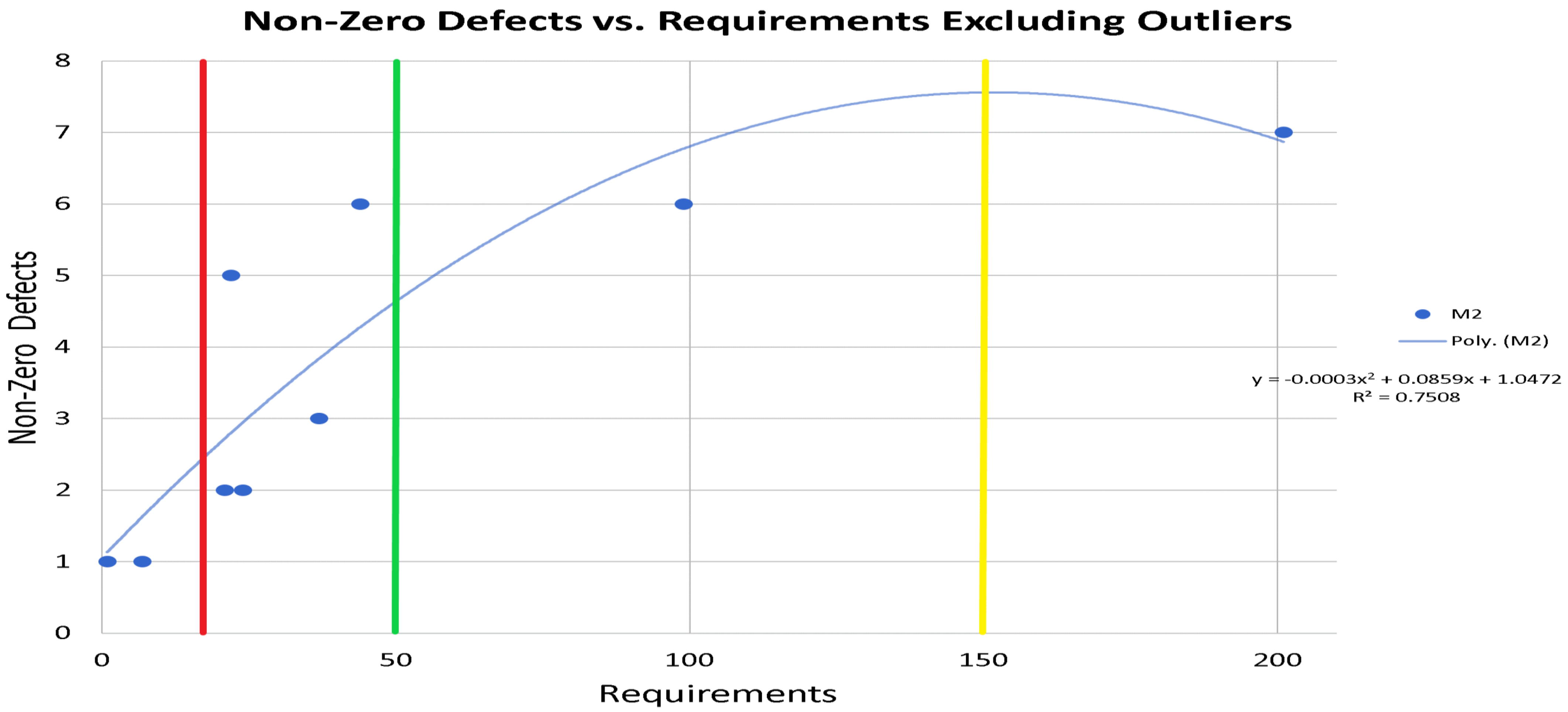
- Safety-critical aerospace system developer
- Highly-regulated industry, ARP4754A, DO-178C, DO-254
- FPR process used by Systems, Software and Firmware engineering



Requirements Review Analysis



Requirements Review Analysis



Conclusions

- FPRs are more effective when conducted early in the life cycle
- There may be a correlation between review size and effectiveness
  - Large review sizes did not yield higher defect detection
  - Small review sizes may not be effective
  - There may be a range of review sizes that optimizes defect detection
- Sample sizes too small to reach definitive conclusions
- Additional studies are necessary to refine model, identify correlations