Robustness Validation and Failure Rates

"Sample Sizes in Reliability Testing"

Werner Kanert – Infineon
René Rongen – NXP

Outline

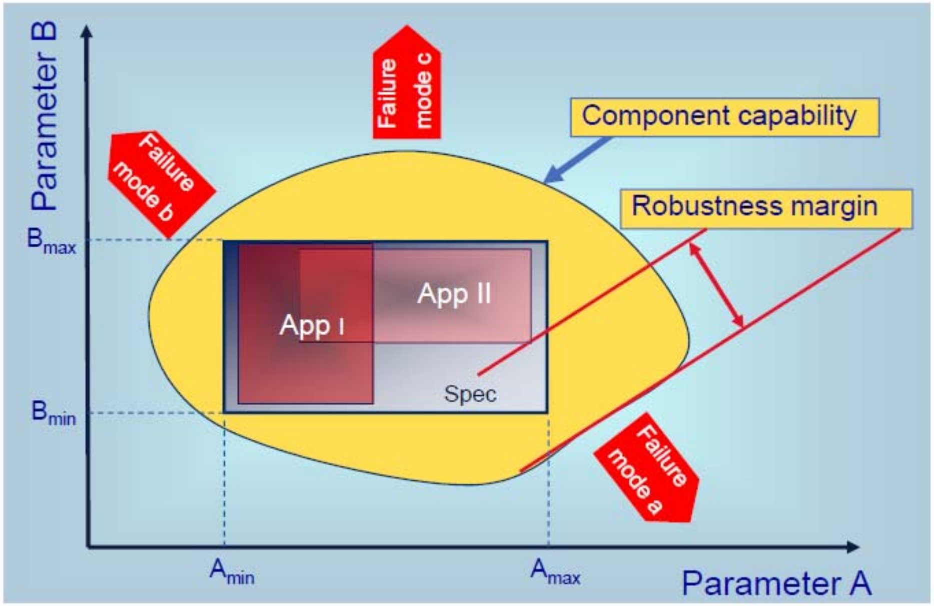
- Robustness Validation: refreshing
- Awareness about numbers in Q100 sampling for reliability
- Defining sample sizes for reliability tests
- Back to Robustness Validation: RIF

Outline

- Robustness Validation: refreshing
- Awareness about numbers in Q100 sampling for reliability
- Defining sample sizes for reliability tests
- Back to Robustness Validation: RIF

What Really Matters…

- Is the device "sufficiently reliable" in the application?
  - Robustness Validation is a method, described in ZVEI handbooks and SAE standards, to demonstrate that a semiconductor component is "sufficiently" reliable



Note, that the component capability can change over life due to degradations

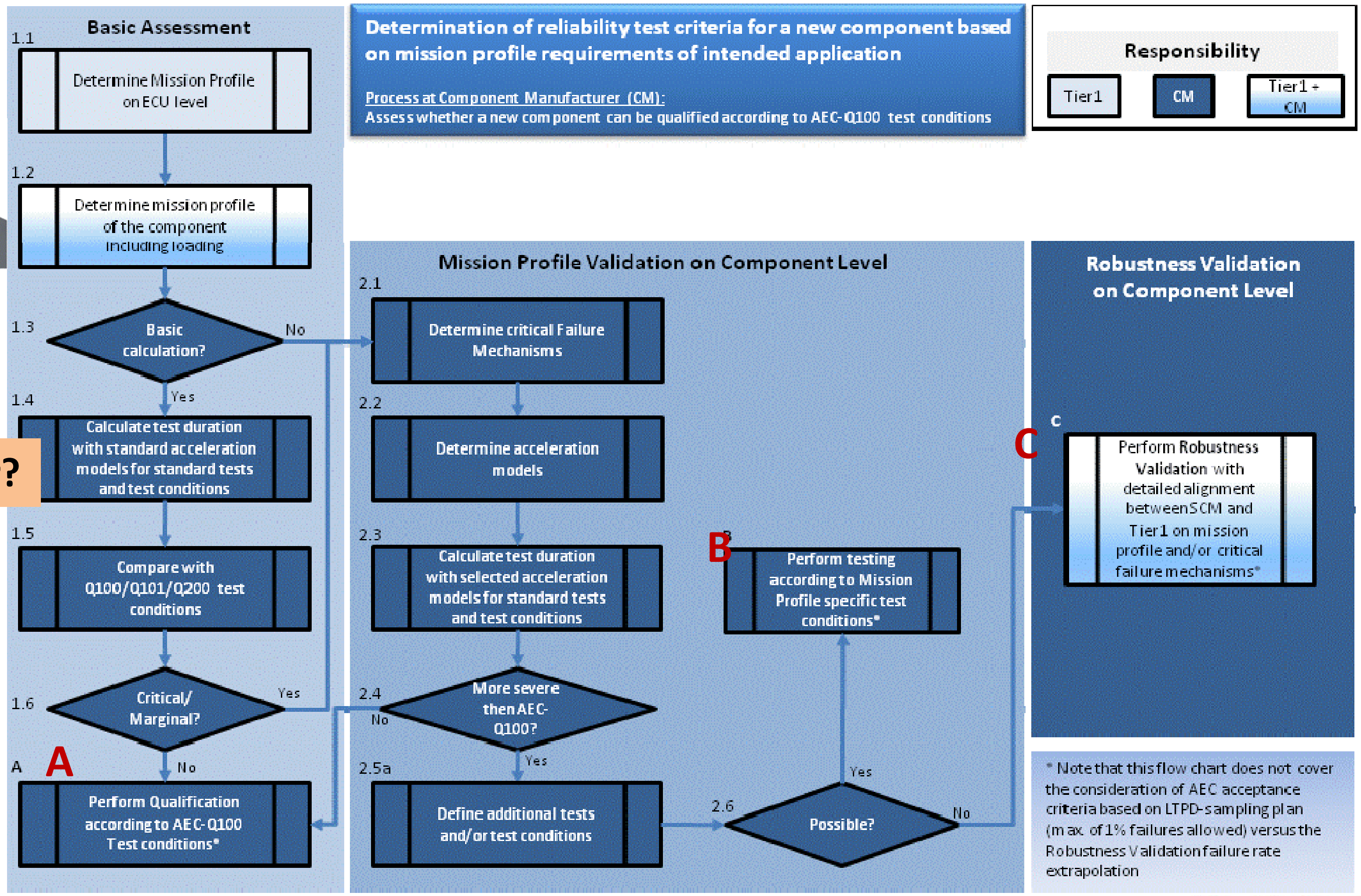
Determination of reliability test criteria for a new component based on mission profile requirements of intended application

Process at Component Manufacturer (CM):
Assess whether a new component can be qualified according to AEC-Q100 test conditions

Automotive Electronics Council
Component Technical Committee

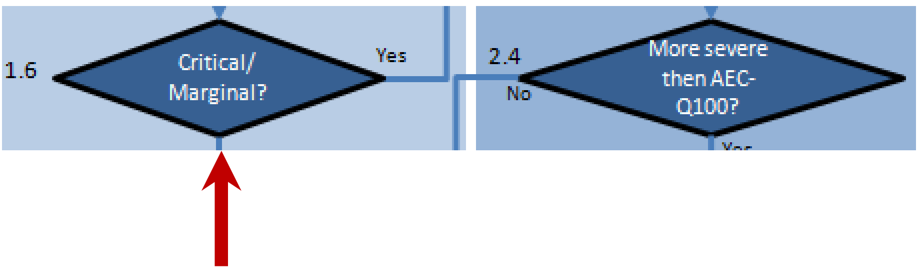
Flow chart 1

Remember?

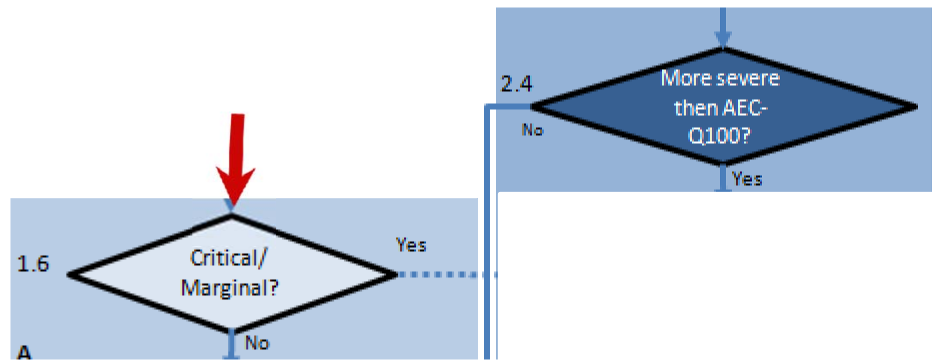


RV Flow Chart (slide from last year)

Flow chart 1



Flow chart 2



Remember?

- The mission profile is translated into an equivalent stress with the same conditions as the qualification standard test. This calculated stress duration t_{calc} (in hours or number of cycles) has to be compared to the standard qualification duration t_{stand}.

But what about failure rates and sample sizes?

Outline

- Robustness Validation: refreshing
- Awareness about numbers in Q100 sampling for reliability
- Defining sample sizes for reliability tests
- Back to Robustness Validation: RIF

LTPD Sampling Plans

Decision on Lot	Population Value (% Defective)	
	≤ Target	≥ Target
Accept	Correct	Type II error (consumer's risk)
Reject	Type I error (producer's risk)	Correct

LTPD (Lot Tolerance Percent Defective):
The LTPD is a defect level that is unacceptable to the consumer. The consumer would like the sampling plan to have a *low probability of accepting* a lot with a defect level as high as the LTPD.

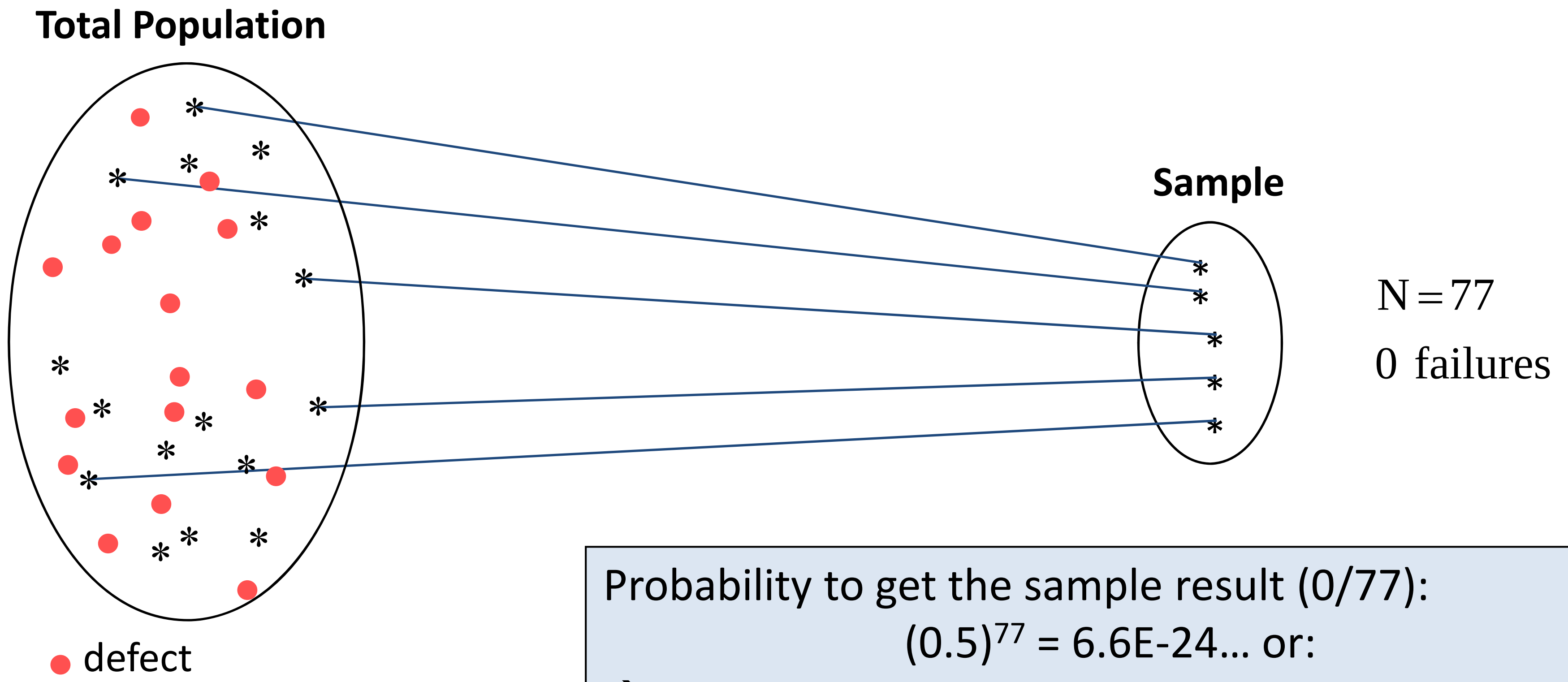
Ref.: MIL-S-19500 and MIL-M-38510

Origin of 0/77

- LTPD plans originated back in the period 1920-40 as part of acceptance sampling.
  - 1/77 = criterion to assure, with 90% confidence, that a lot having 5% defective is rejected
  - However … to allow fails as part of your qualification requirements is not good for the image
  - Therefore, zero fail tests (success-runs) were adopted for qualification

Statistical Meaning of 0/77 (Exercise 1)

Question 1: How likely is this result, when actual defectivity is 50% (500.000 PPM)?

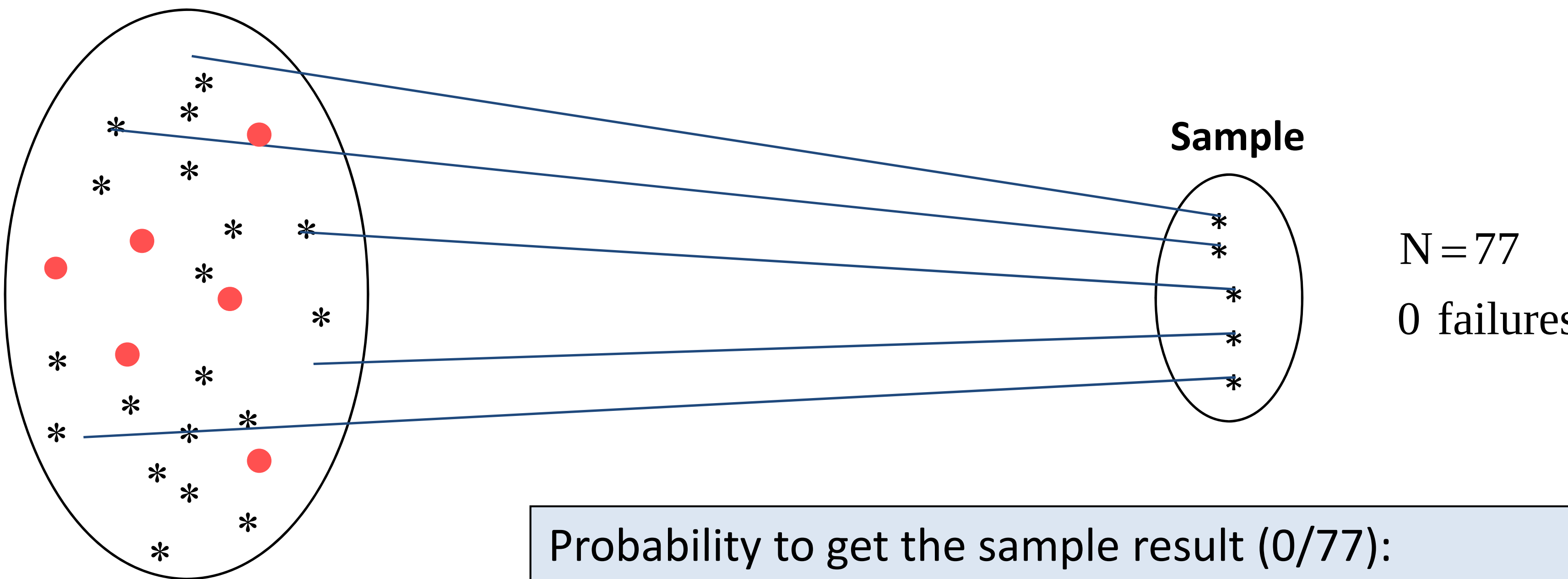


Probability to get the sample result (0/77):
$(0.5)^{77}$ = 6.6E-24… or:
→ With almost 100% confidence: "PPM < 500.000"

Statistical Meaning of 0/77 (Exercise 2)

Question 2: How likely is this result, when actual defectivity is 5% (50.000 PPM)?

Total Population



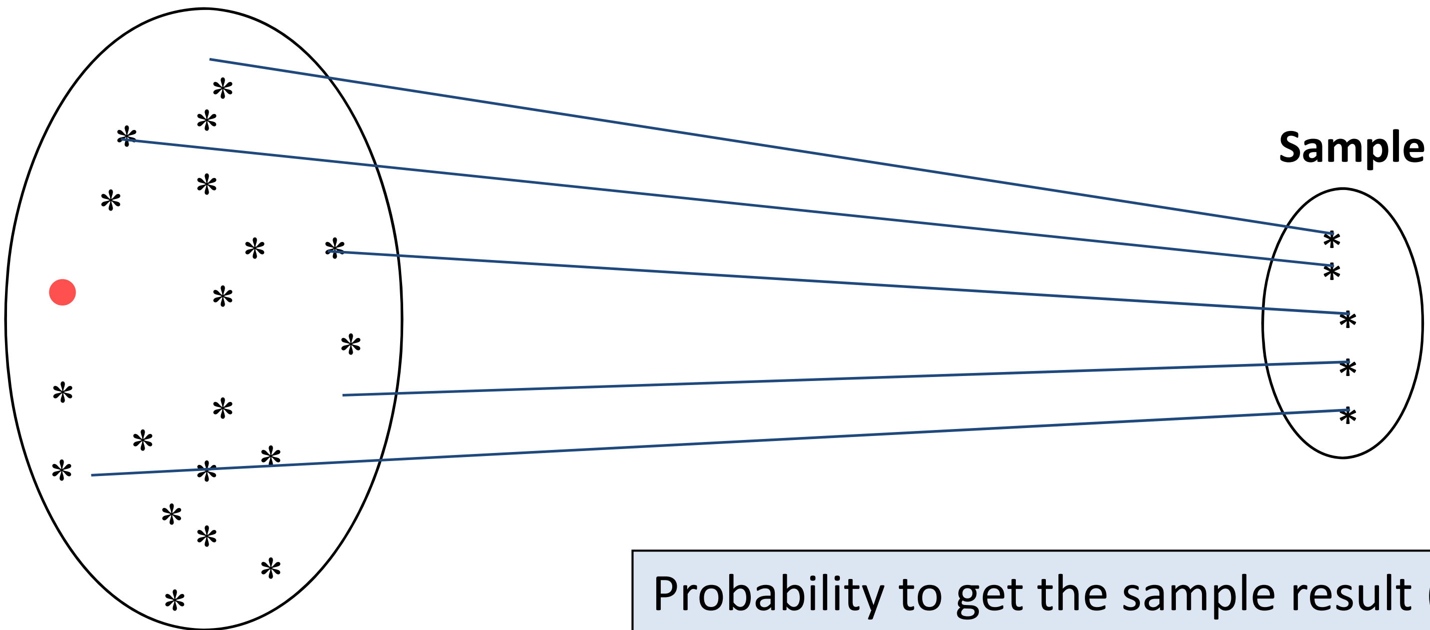
● defect

Probability to get the sample result (0/77):
$(0.95)^{77} = 0.019\ldots$ or:
→ With 98.1% confidence: "estimated PPM < 50.000"

Statistical Meaning of 0/77 (Exercise 3)

Question 3: How likely is this result, when actual defectivity is 3% (30.000 PPM)?

Total Population



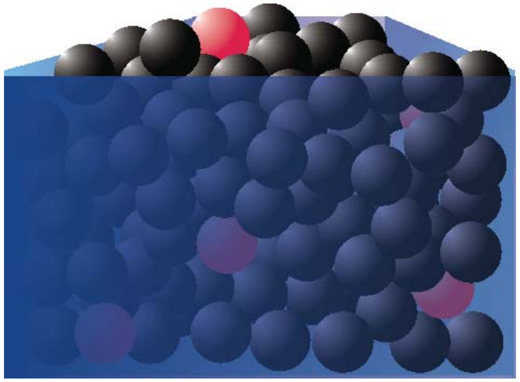
N = 77
0 failures

● defect

Probability to get the sample result (0/77):
$(0.97)^{77} = 0.10...$ or:
→ With 90% confidence: "estimated PPM < 30.000"

Statistical Meaning of 0/77 (Summary)

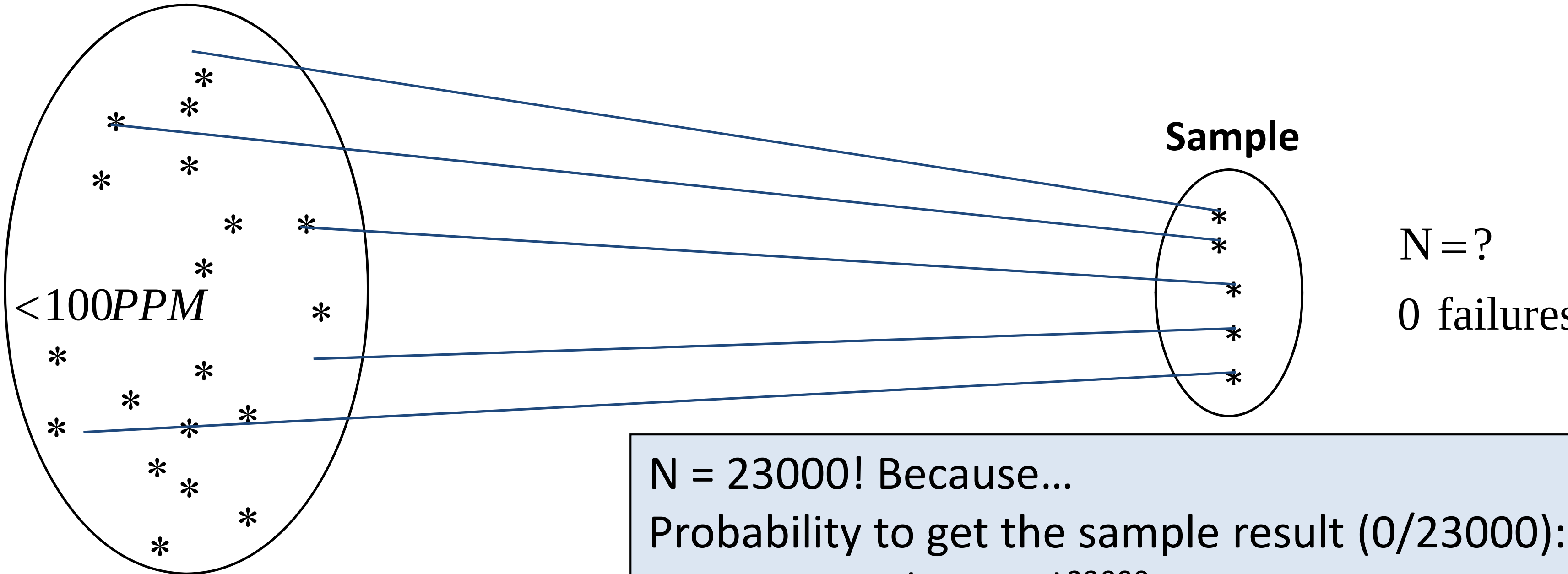
- With a result of 0 failures for a small sample of 77 devices from a much larger total population of which the actual defectivity level is unknown, this means:
  - By lowering the confidence level the prediction is more optimistic
  - By increasing the confidence level the prediction is more pessimistic



Consequence for 100 PPM with 90 % Confidence

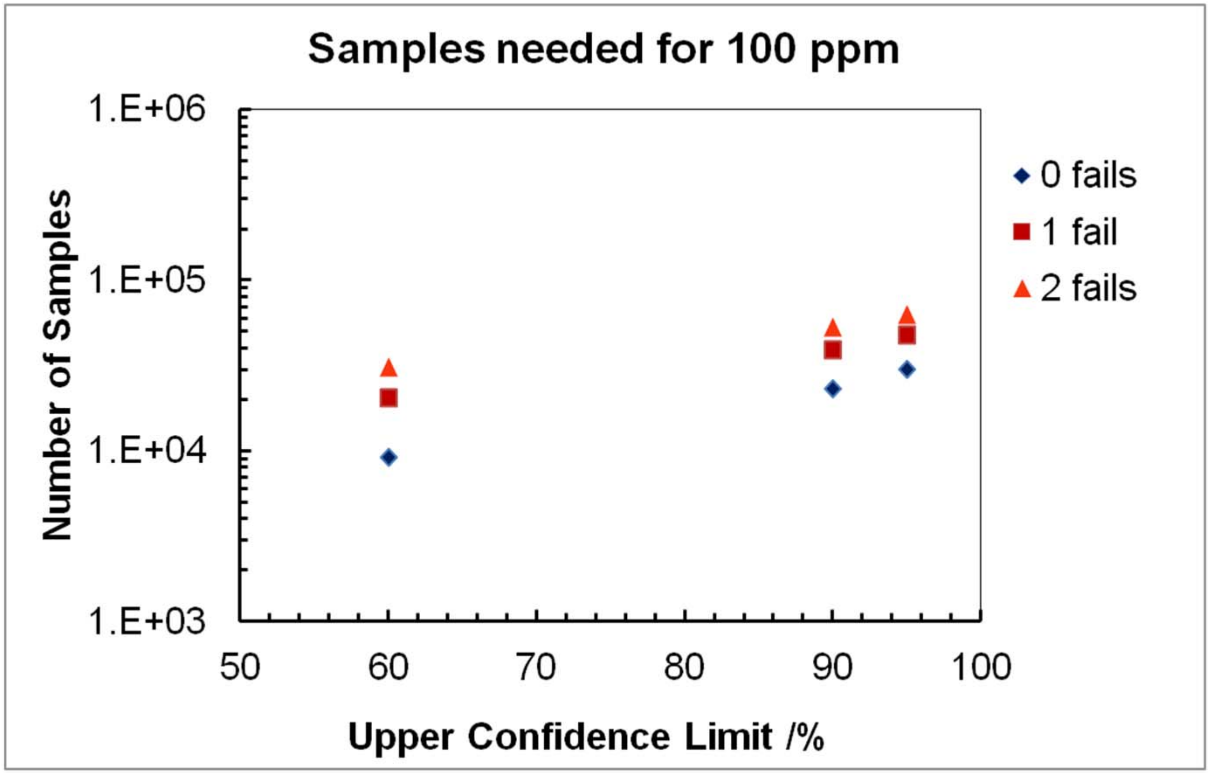
Question 4: How large should a sample be to give an estimated PPM < 100 assuming no fails after test?

Total Population



N = 23000! Because…
Probability to get the sample result (0/23000):
$(0.9999)^{23000} = 0.10$… or:
→ With 90% confidence: "estimated PPM < 100"

Sample Size Needed to Prove 100 ppm



Proving low level failure probabilities by simply accumulating statistics is no viable approach for qualification.

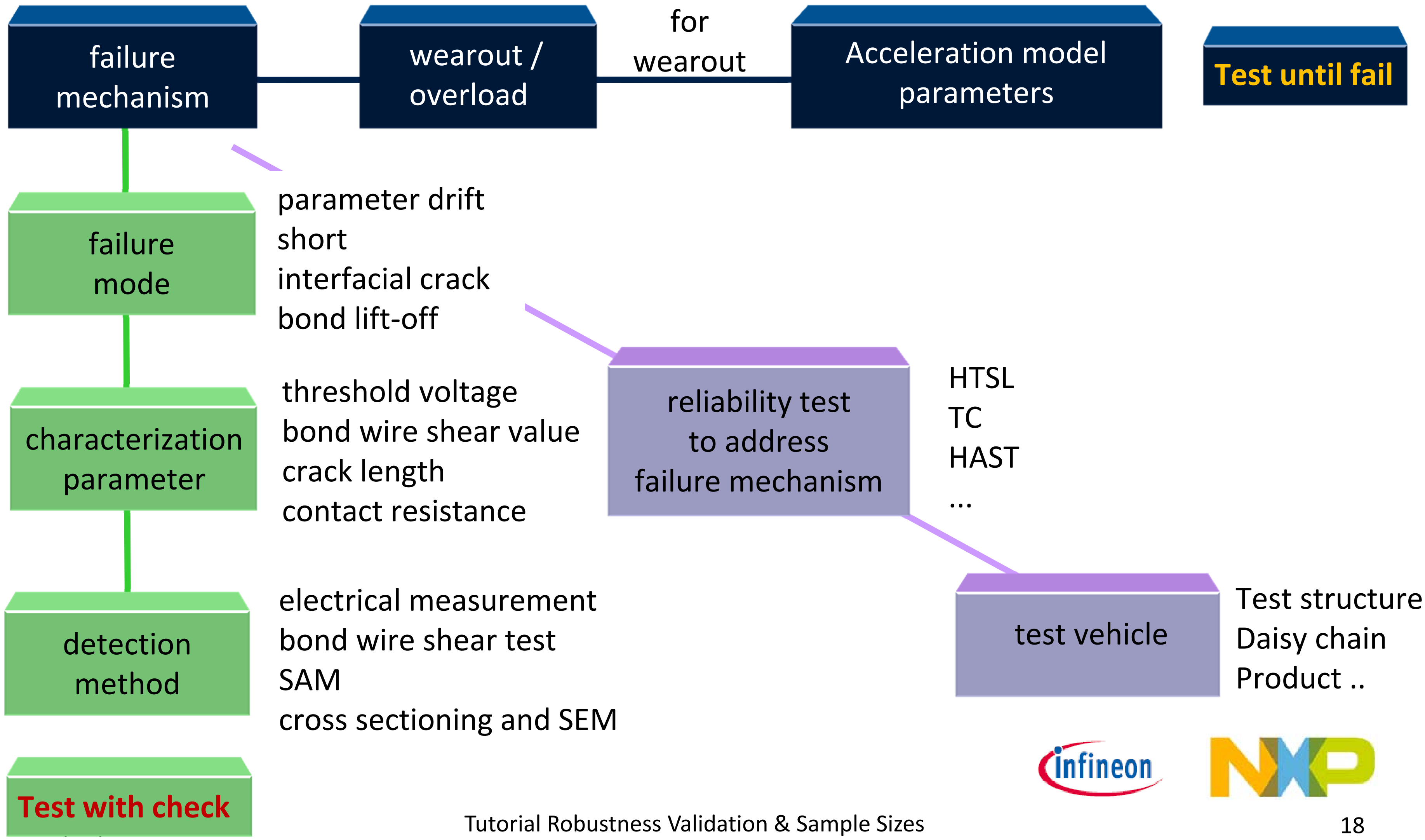
Outline

- Robustness Validation: refreshing
- Awareness about numbers in Q100 sampling for reliability
- Defining sample sizes for reliability tests
- Back to Robustness Validation: RIF

Defining Sample Sizes for Reliability Tests

Pass / fail decision based on fulfillment of electrical spec may not be sufficient and 0/45 would be almost as good as 0/77

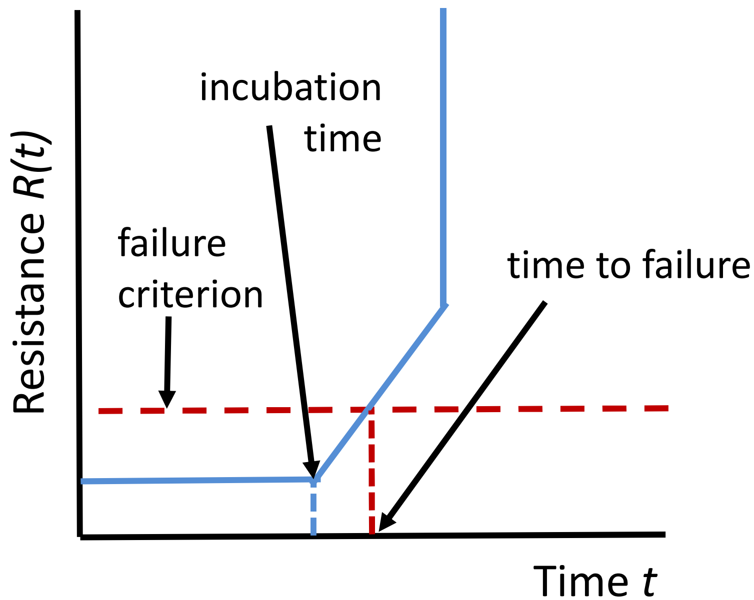
Setting up a Qualification Test

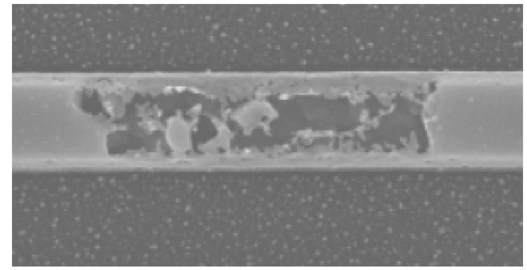


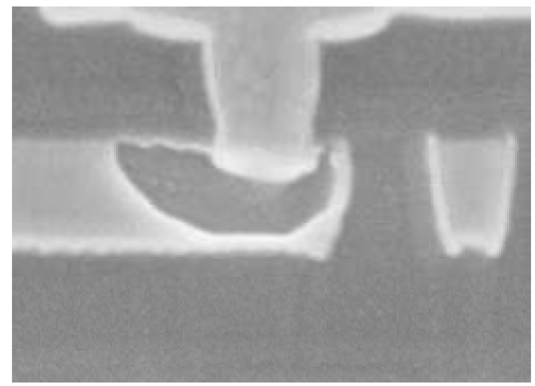
Example (1): Electromigration

Test until fail

Assess degradation by monitoring resistance



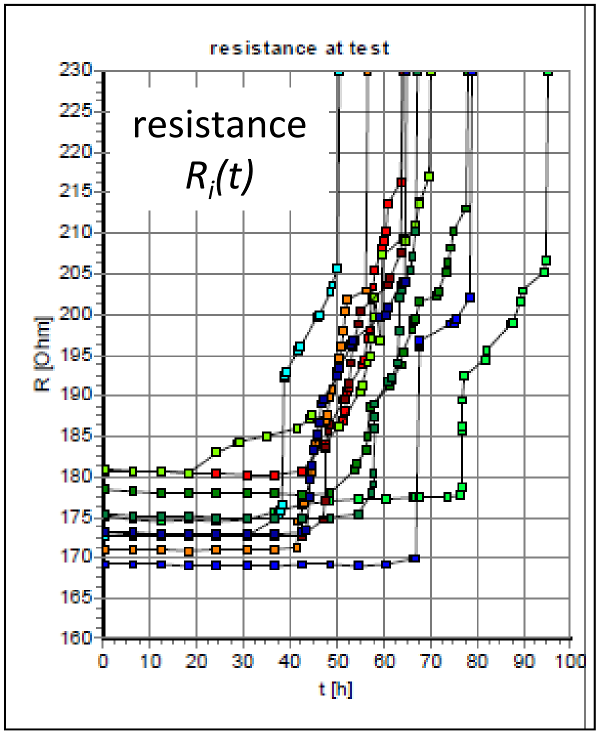




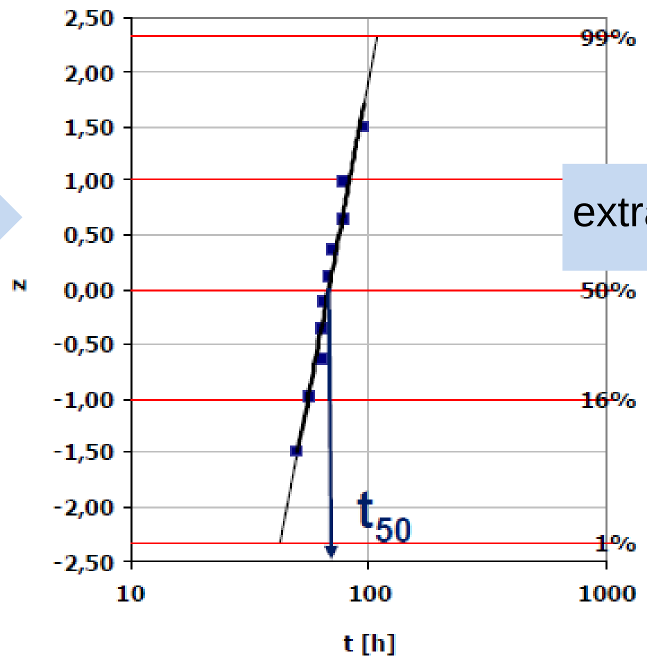
Time To Failure (TTF) is defined by the time until a certain resistance increase to original value $(\Delta R/R_0)$ is reached.

Example (1): Electromigration

Test until fail



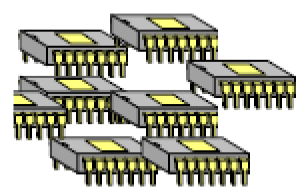
time to failure



extract

t_{50}
σ

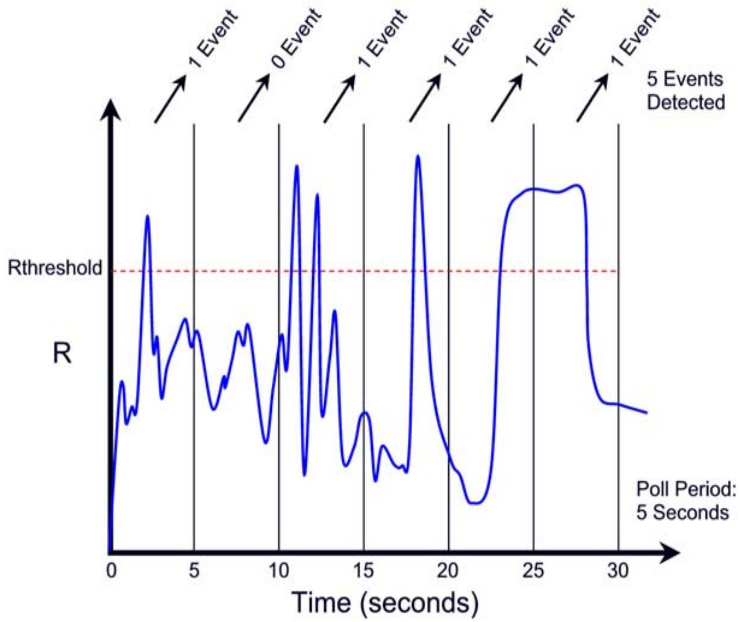
sample size
10-20 per test

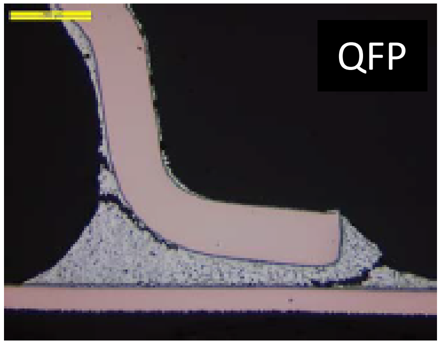


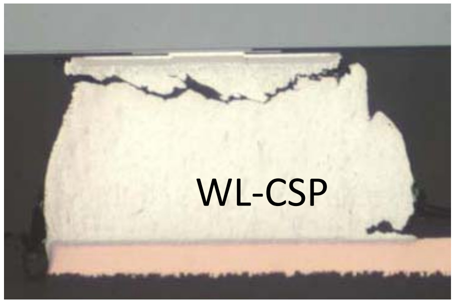
Example (2): Solder Joint Fatigue

Test until fail

Assess degradation by “event detection”



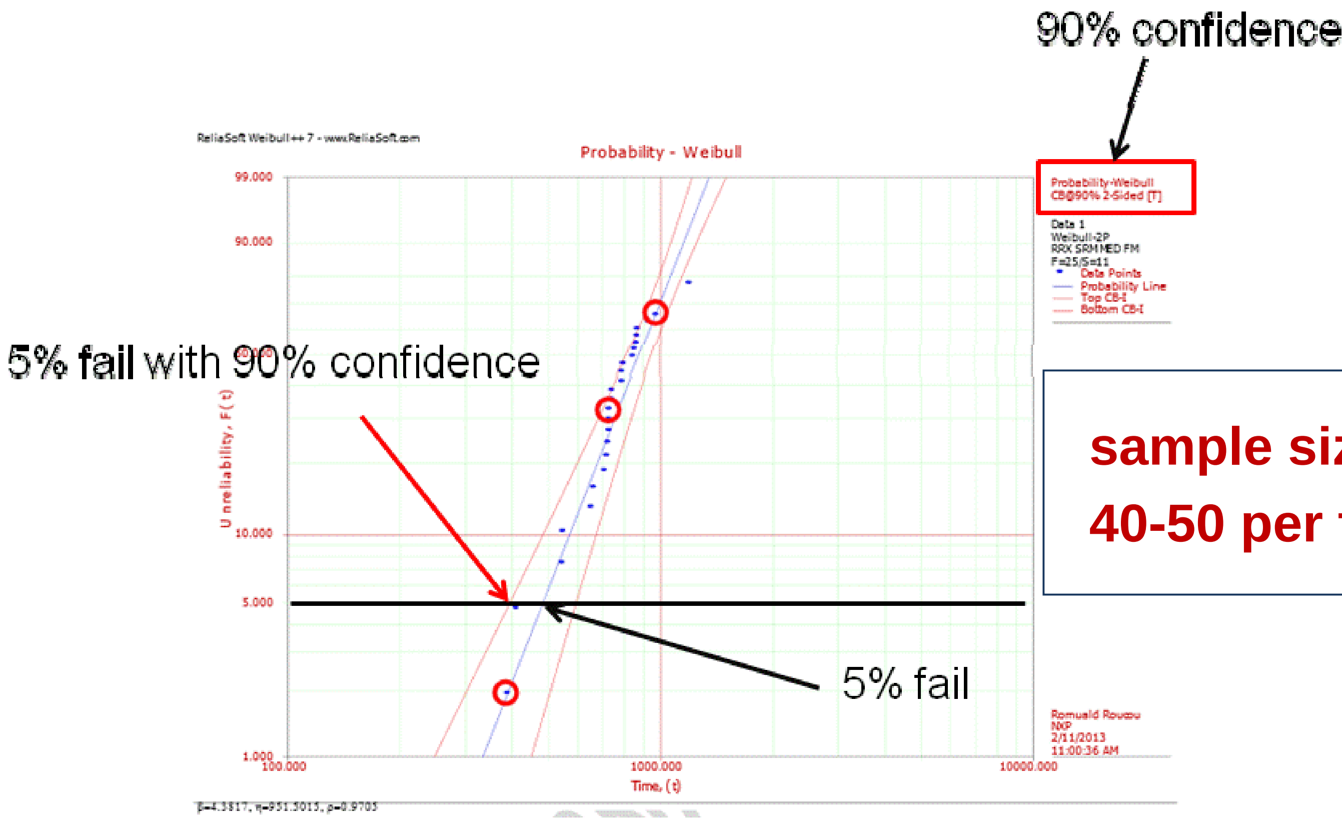




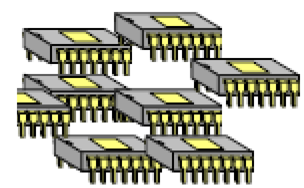
Time To Failure (TTF) is defined by the maximum number of sequential events in which a certain resistance increase to original value ($\Delta R/R_0$) is observed.

Example (2): Solder Joint Fatigue

Test until fail



sample size
40-50 per test



Example (3): Wire Bond Fails

Test with check

failure mechanism	thermo-mechanical fatigue of wire connections to chip or lead (nailhead/wedge)
failure mode	bond lift-off break of wire break of wedge
characterization parameter	bond pull values bond shear values wedge pull values
detection method	bond pull test bond shear test wedge pull test

reliability test	test vehicle
temperature cycling	32 pin package

Example (3): Wire Bond Fails

Test with check

Temperature cycling

readouts	characterization	samples
initial	WBP WBS WP add SS	5 parts/32 wires each 5 parts/32 wires each 5 parts/32 wires each 5 parts
1000 cycles	WBP WBS WP add SS	5 parts/32 wires each 5 parts/32 wires each 5 parts/32 wires each 5 parts
2000 cycles	WBP WBS WP add SS	5 parts/32 wires each 5 parts/32 wires each 5 parts/32 wires each 5 parts
3000 cycles	WBP WBS WP add SS	5 parts/32 wires each 5 parts/32 wires each 5 parts/32 wires each 5 parts

sample size
80 per test

calculate backward how many parts are needed at the start

Does not take into account lot to lot variations and process corners

WBP – wire bond pull test
WBS – wire bond shear test
WP – wedge pull test
SS – suspended samples

Example (3): Wire Bond Fails

Test with check

- Failure mechanism: Thermo-mechanical fatigue
- Failure mode: Bond lift-off, wedge break
- What do I want to look at (characterization parameter)?
  - Degradation of bond pull value, bond shear value, wedge pull value
- Define duration of test and intermediate readouts
- How many devices are needed per readout?
- How many devices should be left over at the end of test?
  E.g. to further extend the test?
- Sum up all devices to get the sample size you start with

Defining Sample Sizes for Reliability Tests (Summary)

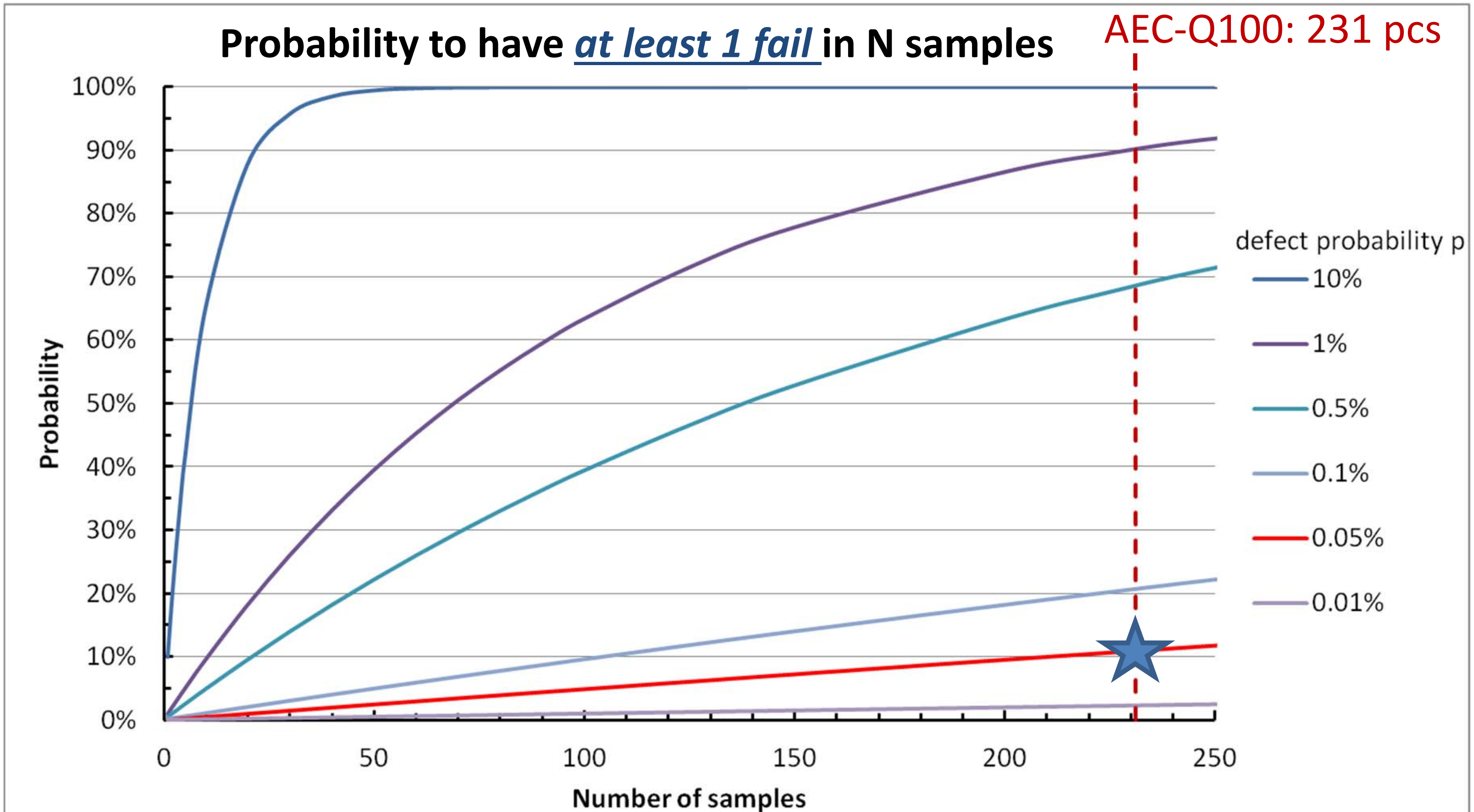
- Different test methods need different approaches, e.g.
  - Accelerated life tests (ALT)
  - Accelerated degradation tests with repeated measurements
  - Accelerated degradation tests with destructive measurements at readouts

- Sample sizes have to be adapted to the requirements determined by the test method

Outline

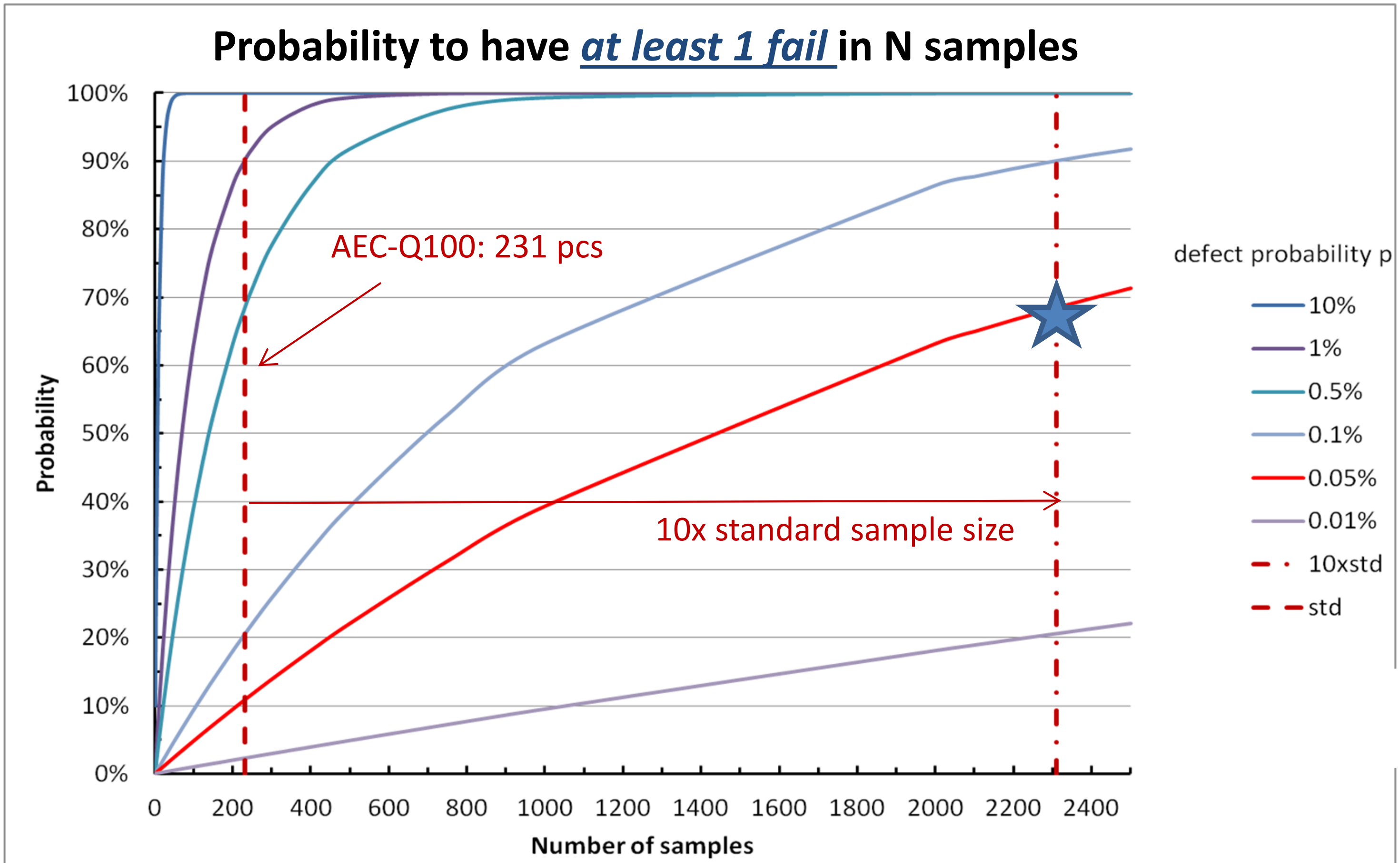
- Robustness Validation: refreshing
- Awareness about numbers in Q100 sampling for reliability
- Defining sample sizes for reliability tests
- Back to Robustness Validation: RIF

Probability of Detecting Failures (1)



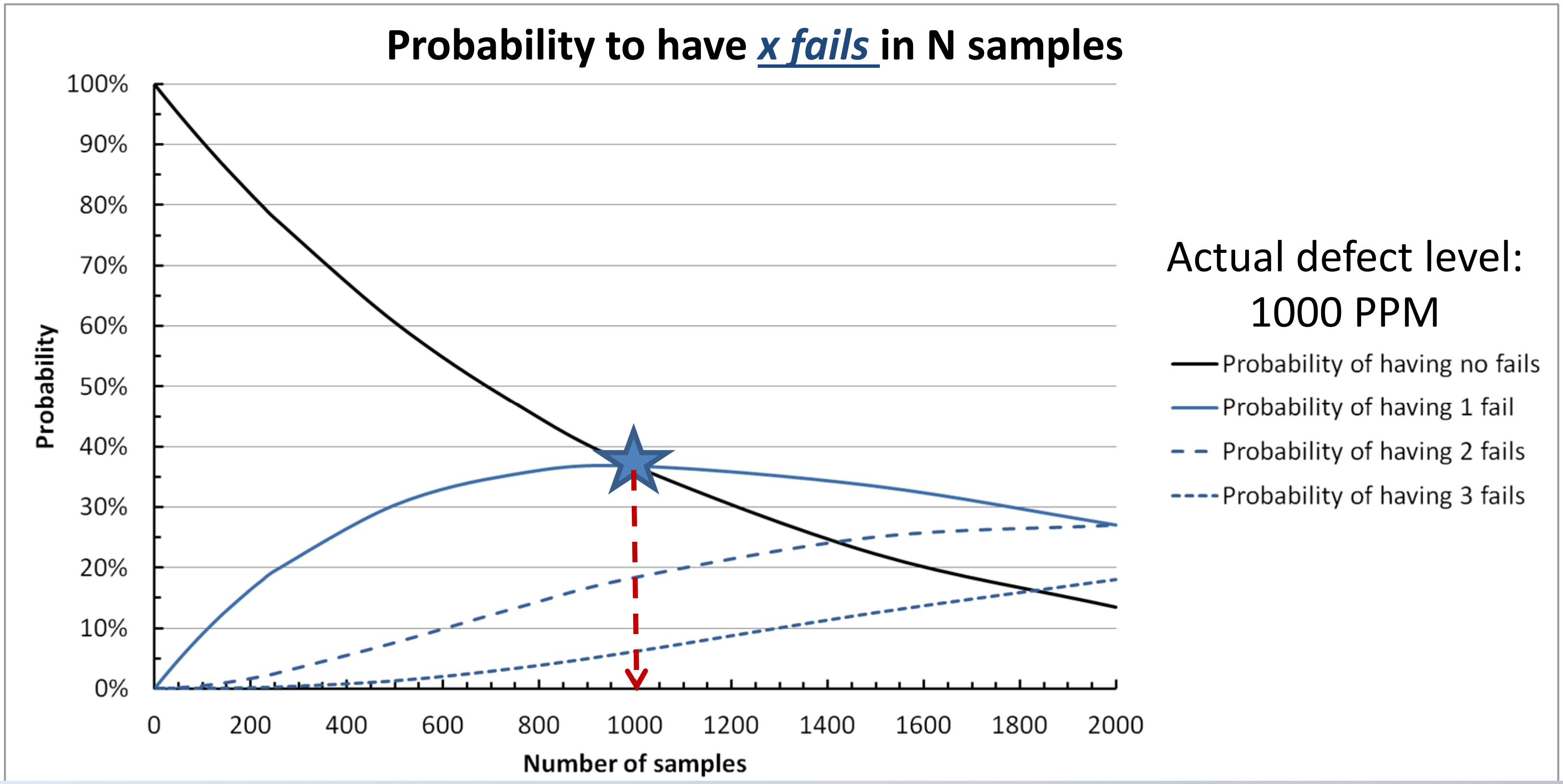
If defect level is 50 PPM: the probability to find a fail is only 10 %

Probability of Detecting Failures (2)



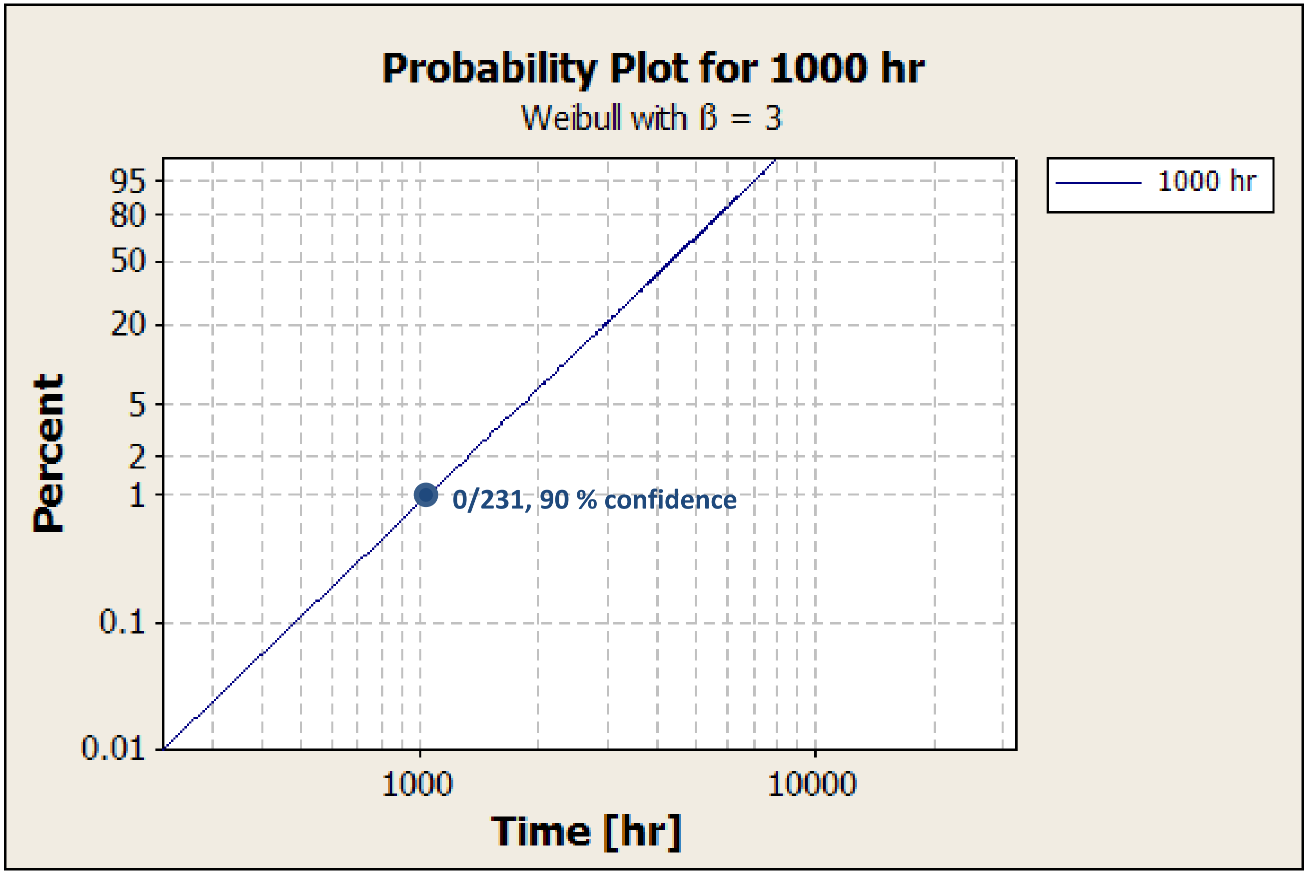
When increasing sample size by a factor of 10: probability <70%, i.e. chance of having no fail is still 1 in 3

Probability of Detecting Failures (3)



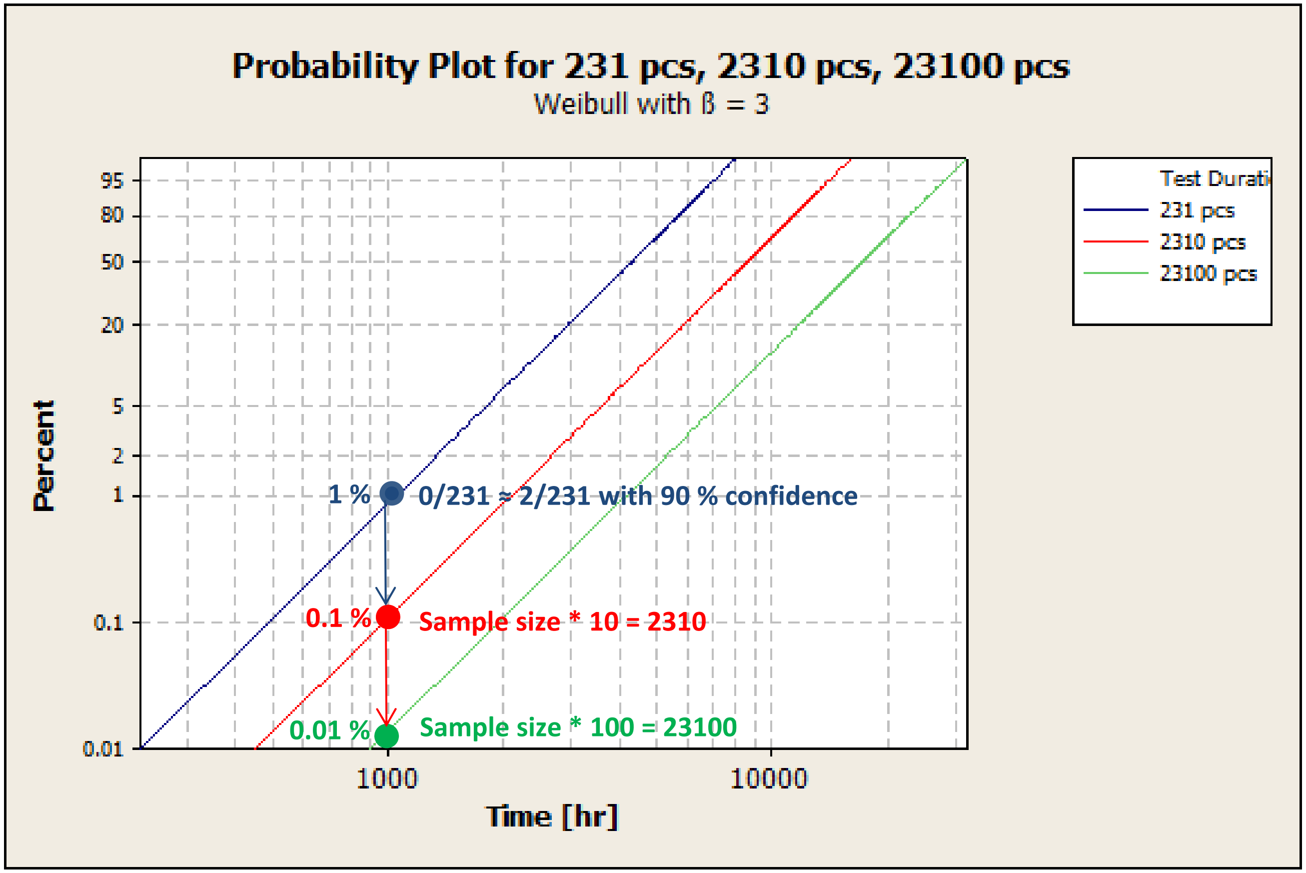
Probability to have 1 fail is maximum around 1000 pcs and is as likely as to have no fail. Increasing the sample size makes no sense.

Weibull with 0 Fails Assuming Wearout (ß = 3)



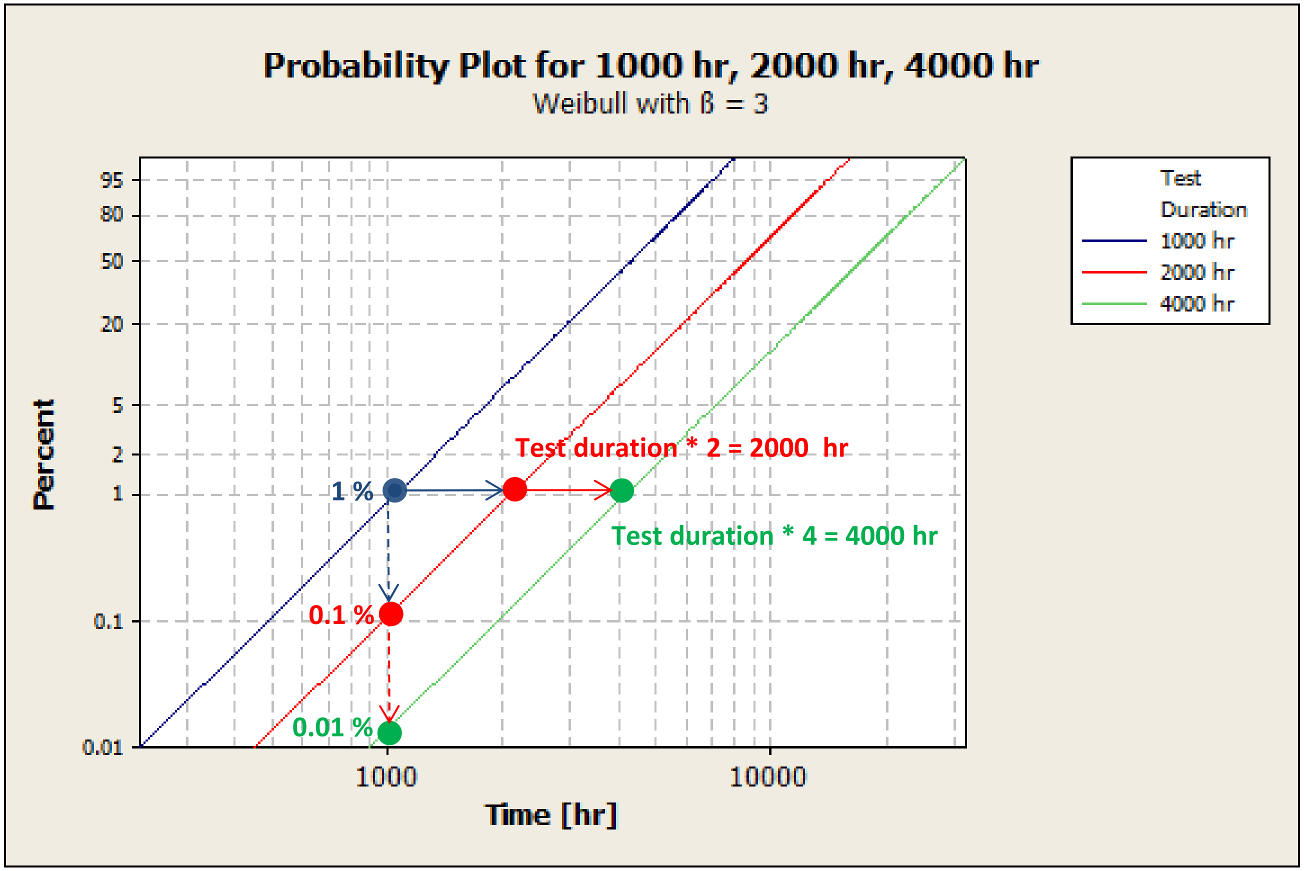
Failure Rate @ 1000 hr: ≈ 1 % (assumed: 0 fails)

Effect of Larger Sample Sizes



By increasing sample sizes: failure rate @ 1000 hr will drop

Effect of Longer Test Times



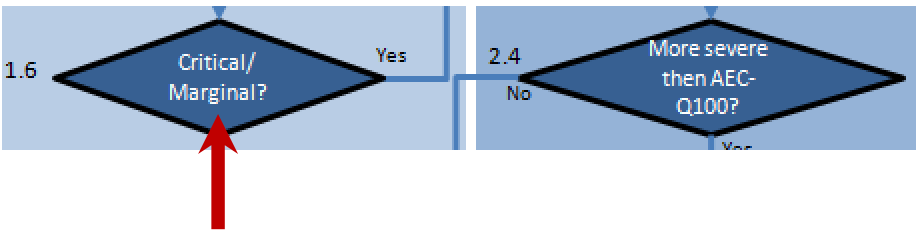
By increasing test duration: failure rate @ 1000 hr will drop

Samples Sizes & Test Times (Summary)

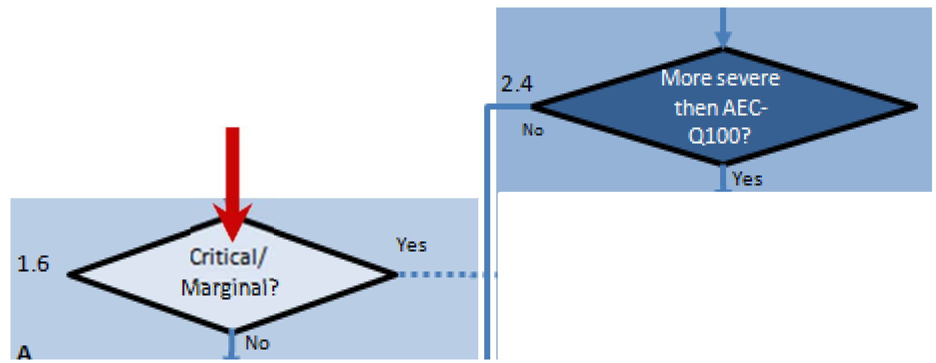
- For a given 1000 h test result of 0/231, the predicted failure rate can be improved by increasing the sample size OR increasing the test time
- Really large sample sizes are needed to significantly improve the statistics, but this increase is limited by the amount of test positions
- However, extended read points allow for determining the (intrinsic) robustness margin, while contributing also to improved statistics at 1000 hr
  - ➡ Robustness Indication Figure (RIF)

RV Flow Chart (slide from last year)

Flow chart 1



Flow chart 2



Remember?

- The mission profile is translated into an equivalent stress with the same conditions as the qualification standard test. This calculated stress duration t_{calc} (in hrs or number of cycles) has to be compared with the standard qualification duration t_{stand}.

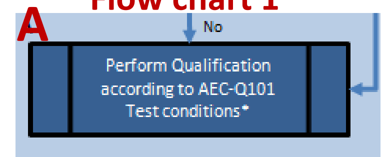
Let us talk about RIF again

Robustness Indicator Figure (RIF)

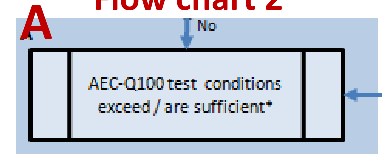
- Compare t_{calc} with t_{stand}:

$$RIF = \frac{t_{stand}}{t_{calc}}$$

Flow chart 1



Flow chart 2



- If RIF ≥ 1, mission profile validated (**A**)
  - Example (from Table A7.1 AEC-Q100):

Humidity (1)	t_u = 131,400 hr (average on/off time over 15 yr of use) RH_u = 74% (average relative humidity in use environment) T_u = 32 °C (average temperature in use environment: 9% 87 °C -time on- and 91% 27 °C –time off)	Temperature Humidity Bias	RH_t = 85 % (relative humidity in test environment) T_t = 85 °C (ambient temperature in test environment)	Hallberg-Peck $A_f = \left(\frac{RH_t}{RH_u}\right)^p \bullet \exp\left[\frac{E_a}{k_B} \bullet \left(\frac{1}{T_u} - \frac{1}{T_t}\right)\right]$ Also applicable for Highly Accelerated Steam Test and Unbiased Humidity Steam Test (Note)	p = 3 (Peck exponent, 3 is to be used for bond pad corrosion) E_a = 0.8 eV (activation energy; 0.8 eV is to be used for bond pad corrosion) k_B = 8.61733 x 10^{-5} eV/K (Boltzmann's Constant)	T_t = 960 hr $t_t = \frac{t_u}{A_f}$

$$RIF = \frac{t_{stand}}{t_{calc}} = \frac{1000}{960} = 1.04$$

- If RIF < 1 (see next slide)

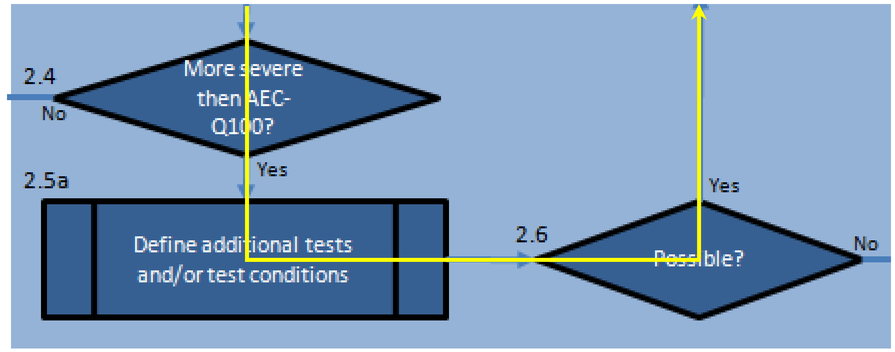
Robustness Indicator Figure (RIF)

- If RIF < 1, then $t_{calc} > t_{stand}$
- Extended test available or possible?
- Compare t_{calc} with $t_{extended}$:

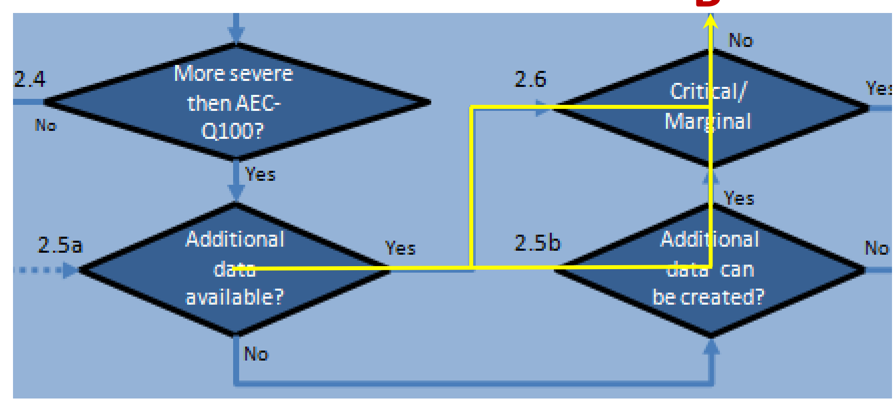
$$RIF = \frac{t_{extended}}{t_{calc}}$$

- RIF ≥ 1, mission profile validated (**B**)
- If RIF < 1 (see next slide)

Flow chart 1



Flow chart 2



Robustness Indicator Figure (RIF)

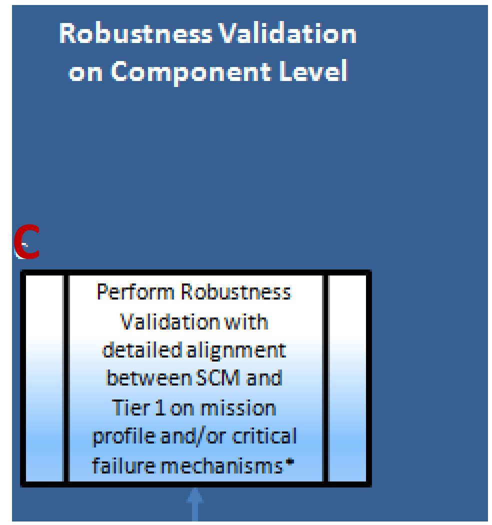
- In case test to fail data are available:

$$RIF = \frac{t_{fail}}{t_{calc}}$$

Examples in Ref. 12

- If RIF ≥ 1, robustness validated (**C**)
- If RIF < 1, assess <u>potential solutions</u>
  (Appendix 7, section A7.3.3.3)
- Lesson 8 and 9:
  - Test-to-fail may result in unrealistic long test duration
  - In these cases change or degradation of electrical or physical properties during or after stress can be used to reduce test duration

Flow chart 1 & 2



Lessons Learned

- 0/77 is rather an industry consensus based value than having a relevant statistical value
- Sample sizes have to be adapted to the requirements determined by the test method
- Testing longer is better than testing more

Questions?