Ureteroscopic treatment of renal stones: Dusting vs Extraction

Khurshid Ghani
Dept of Urology
Will Meurer
Depts of Emergency Medicine and Neurology
University of Michigan

Background

- Use of flexible ureteroscopy (URS) for treating patients with kidney stones has increased compared with other methods
- Technique: Stones are broken down with a laser fiber and are either (a) broken into small fragments (dust) which are left in place to drain spontaneously or (b) broken down and then extracted using baskets and devices.
- Technique (b) costs more money and takes longer time
- There is no consensus among urologists as to whether

Knowledge Gap

- There is no consensus among urologists about which technique results in a better stone-free rate.
- Studies looking at this show equivalent results, yet many urologists continue to extract (this can have complications)

RCT between dusting vs extraction

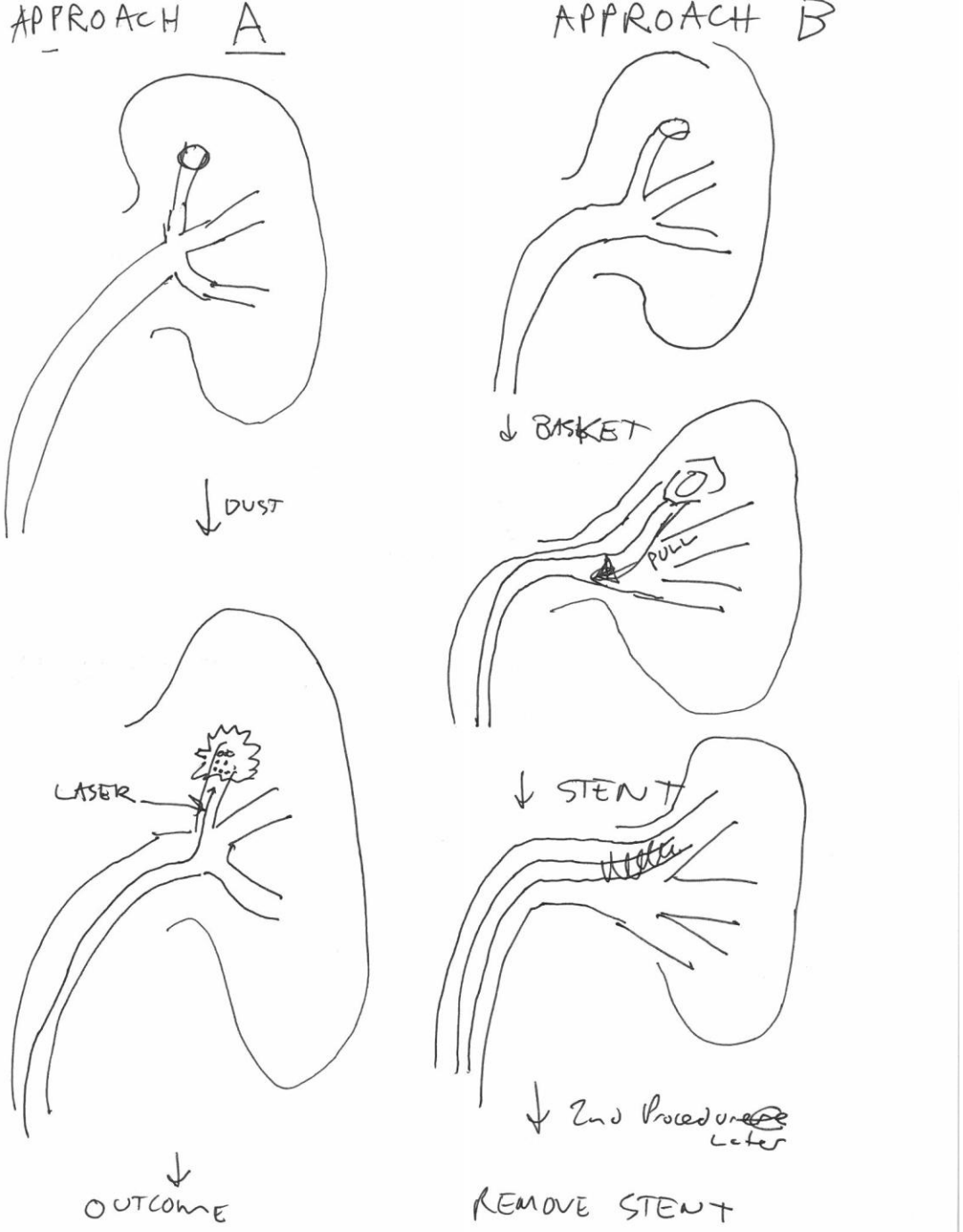
- Hypothesis: stone-free rate after extraction is 70%, and that after dusting is 50%.
- Maybe there are specific stone sizes that are most suited to extraction or dusting?
- Maybe there are specific stone hardness that are more suited to one method or the other?
- Can an adaptive RCT help us answer this question?

Outcome Measure

- Primary: CT scanning : Stone Free Rate = SFR
- Secondary: How many procedures required/cost?

Main covariate

- Stone Size
- Probably range 5mm – 20 mm
- Smaller <5 – everyone would “dust”
- Larger > 20 – everyone would use multi-stage procedure


APPROACH A
APPROACH B
DUST
BASKET
PULL
LASER
STENT
2nd Procedure Later
OUTCOME
REMOVE STENT

Trial Requirements

- Approximately 100-200 subjects
- Primary outcome assessments radiological (could be blinded)
- Comparative effectiveness (both treatments reasonable

Main question to answer

- Who should get what?
- Where is transition between small enough to "dust" and "clearly do two-step procedure"
- Can you find non-inferiority of "dust" (the cheaper way) under some threshold?
- Can you find superiority of "two-step" over some threshold?
- Might you find some middle space where you clearly won't separate (enrich?)

Proposed Adaptive Design

- Enroll N1 patients 5-20mm stones, equally randomized
- Perform logistic regression fitting

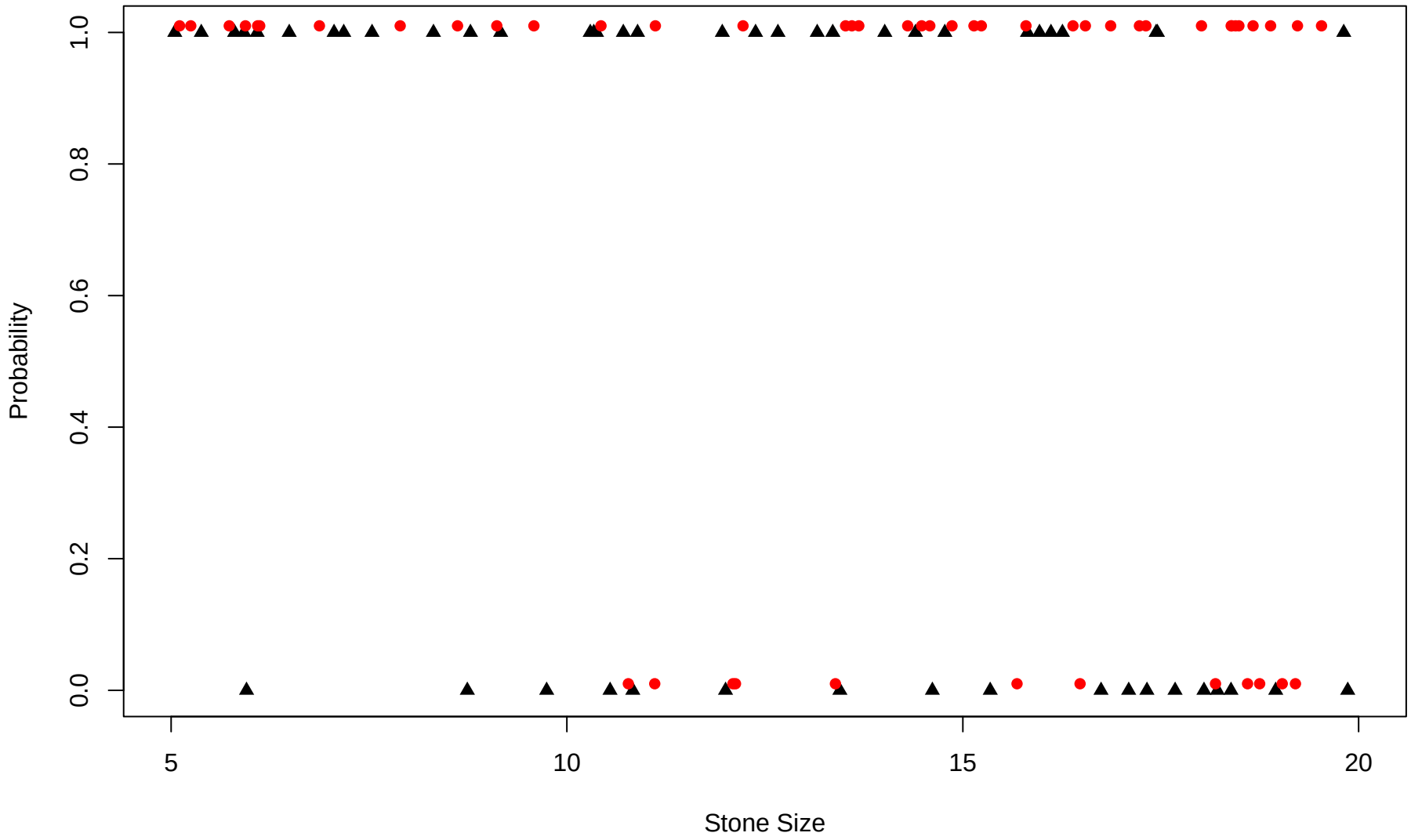
  Pr(Success) ~ Stone size*Group
- Superiority: For each stone size calculate

  Pr(Sup)=Pr(Pr(Success | 2step) > Pr(Success | Dusting))
- Non-inferiority: For each stone size calculate

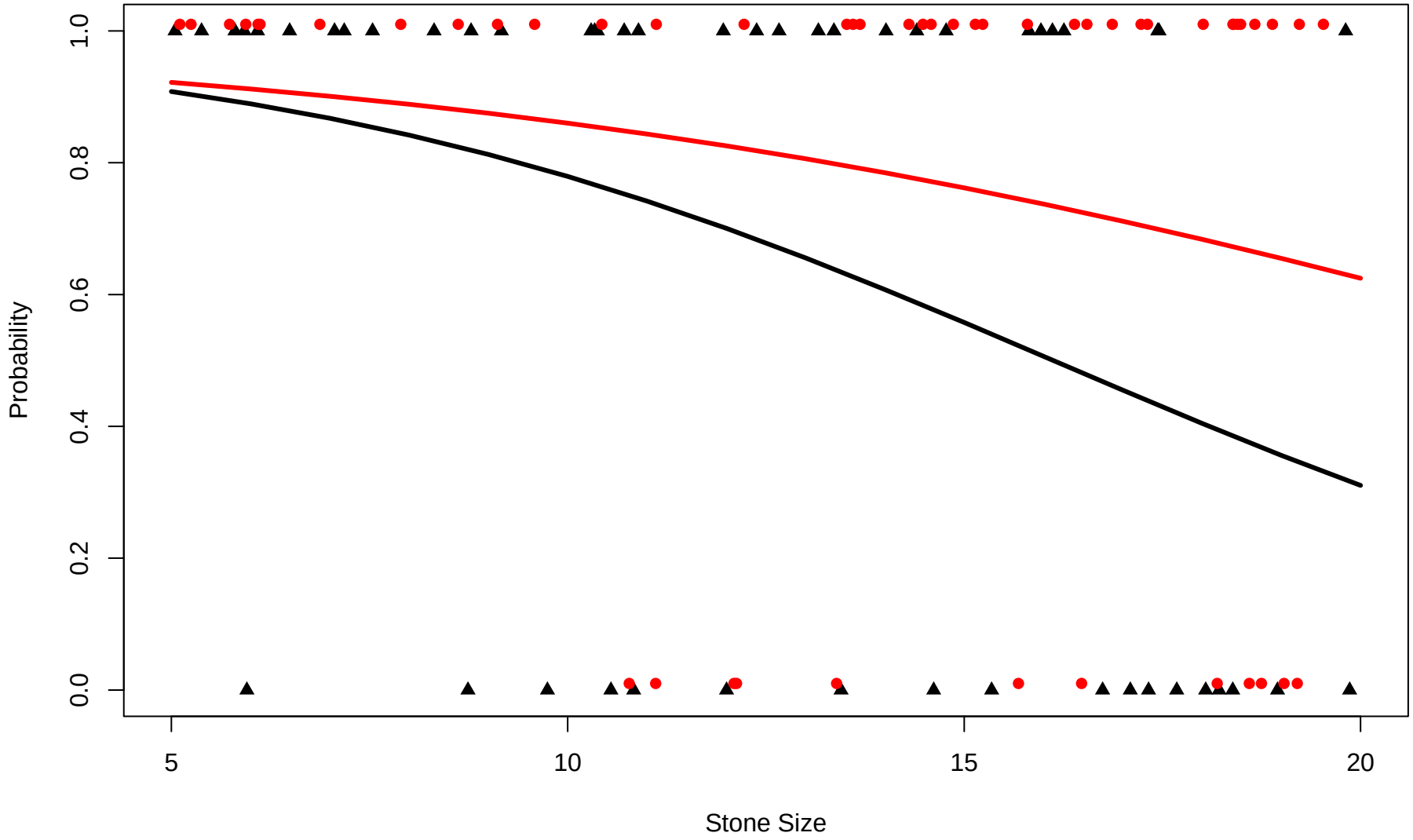
  Pr(NI)=Pr(Pr(Success | Dusting) – Pr(Success | 2step) > -0.10)
- Stop enrolling sizes for which Pr(2 step Sup)>0.95

  or Pr(Dusting NI) > 0.95
- Enroll N2 more patients with stone sizes for which superiority and non-inferiority isn't known.

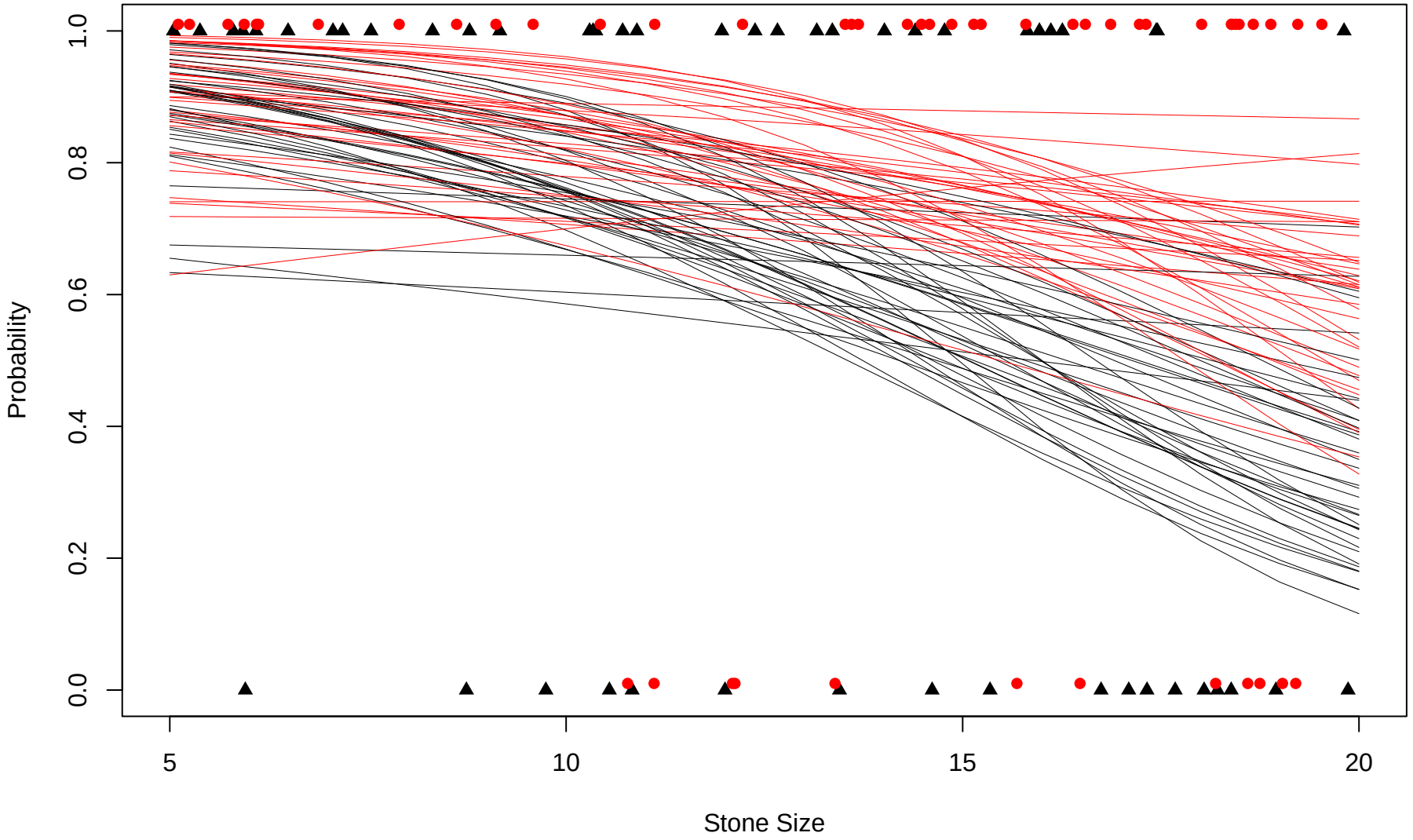
Example Trial, N1=N2=100



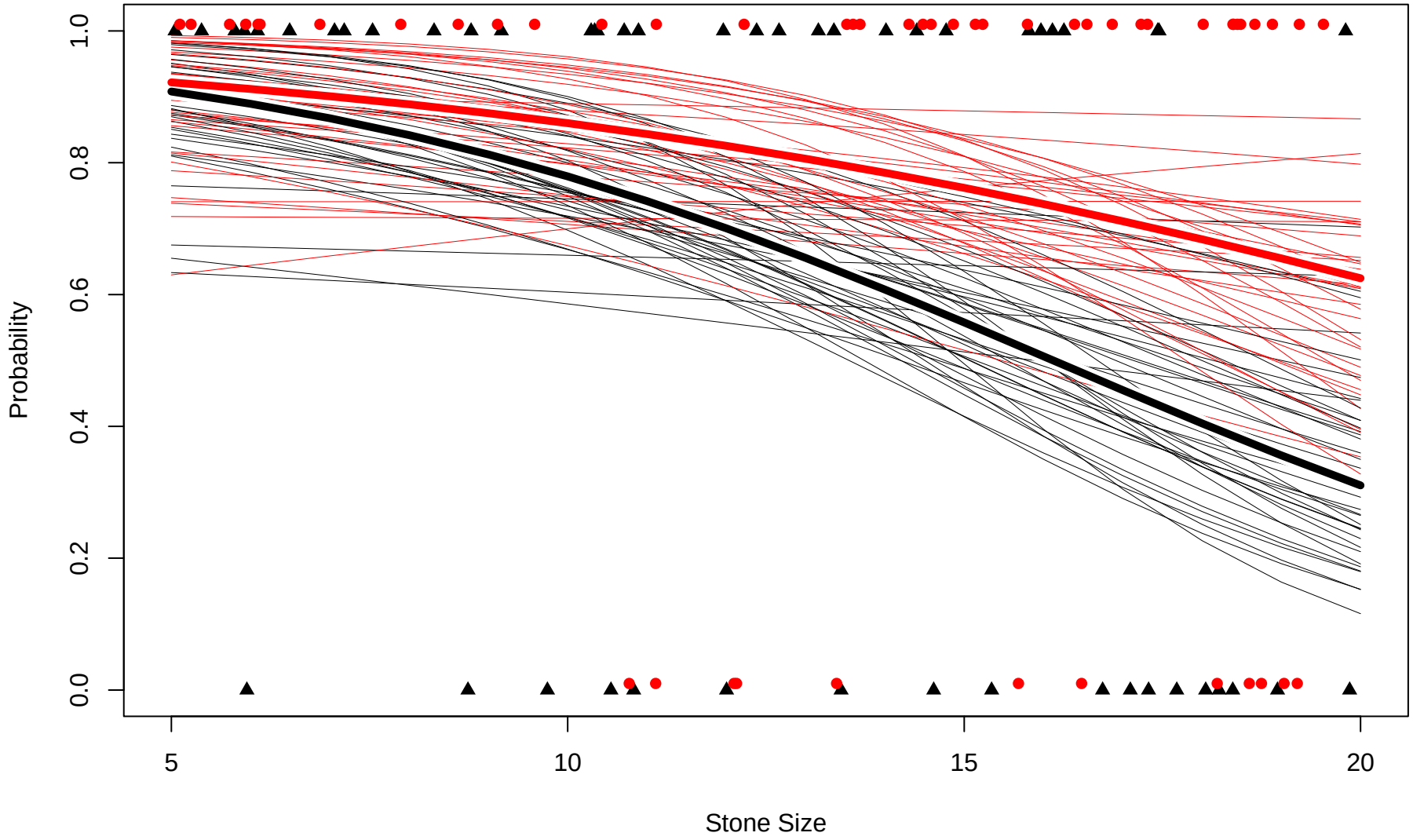
Ex1: Interim Analysis after N1=100



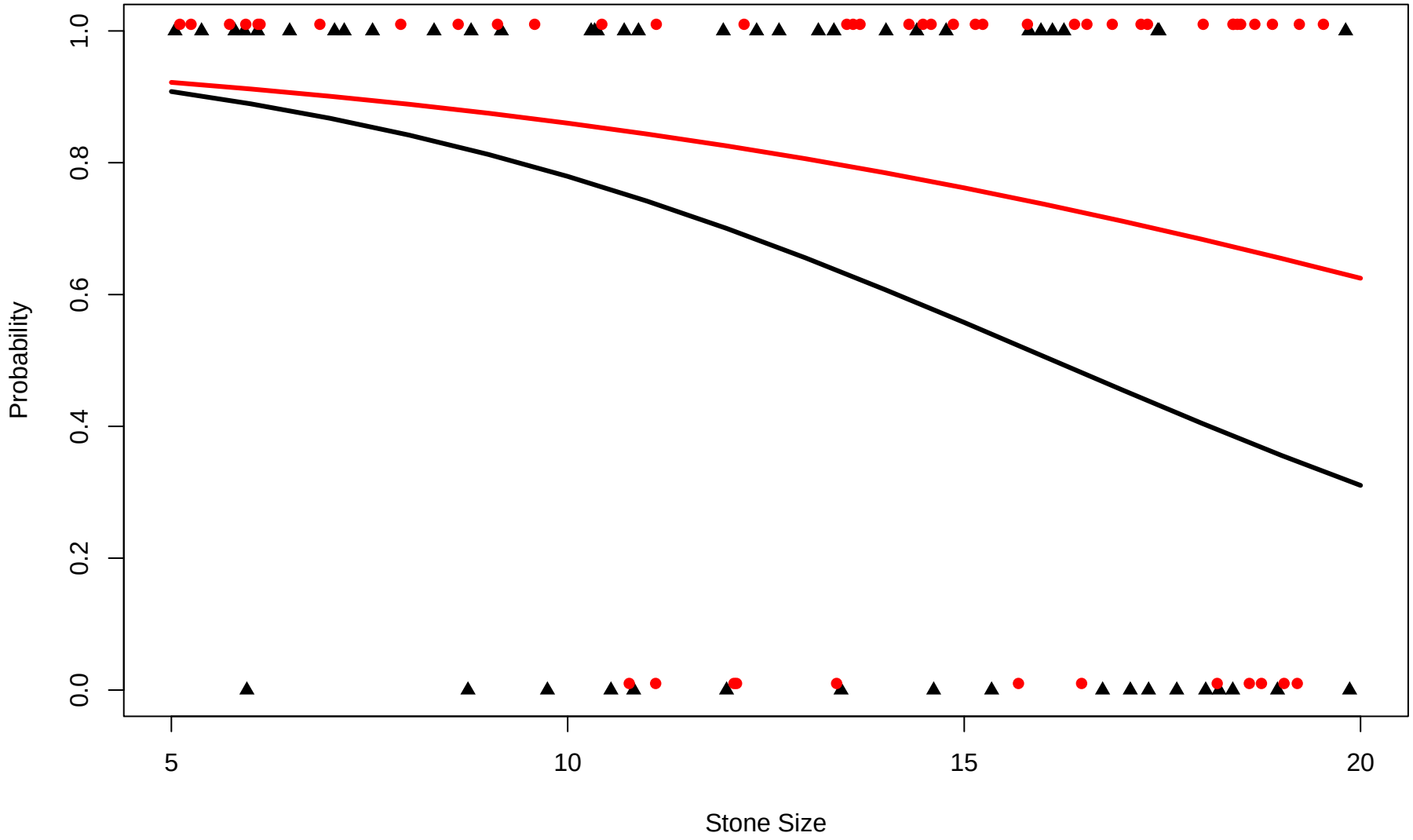
Ex1: Interim Analysis after N1=100



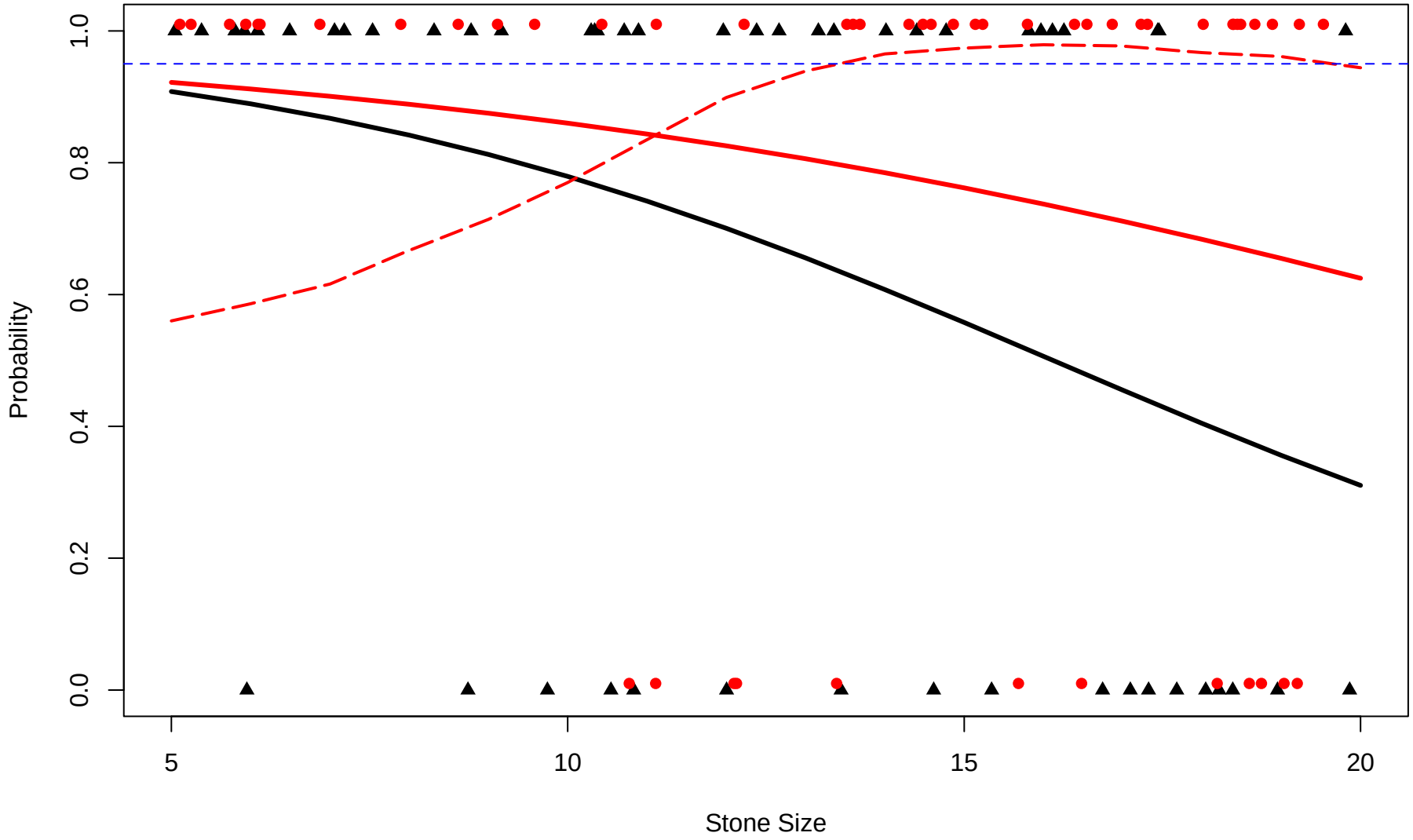
Ex1: Interim Analysis after N1=100



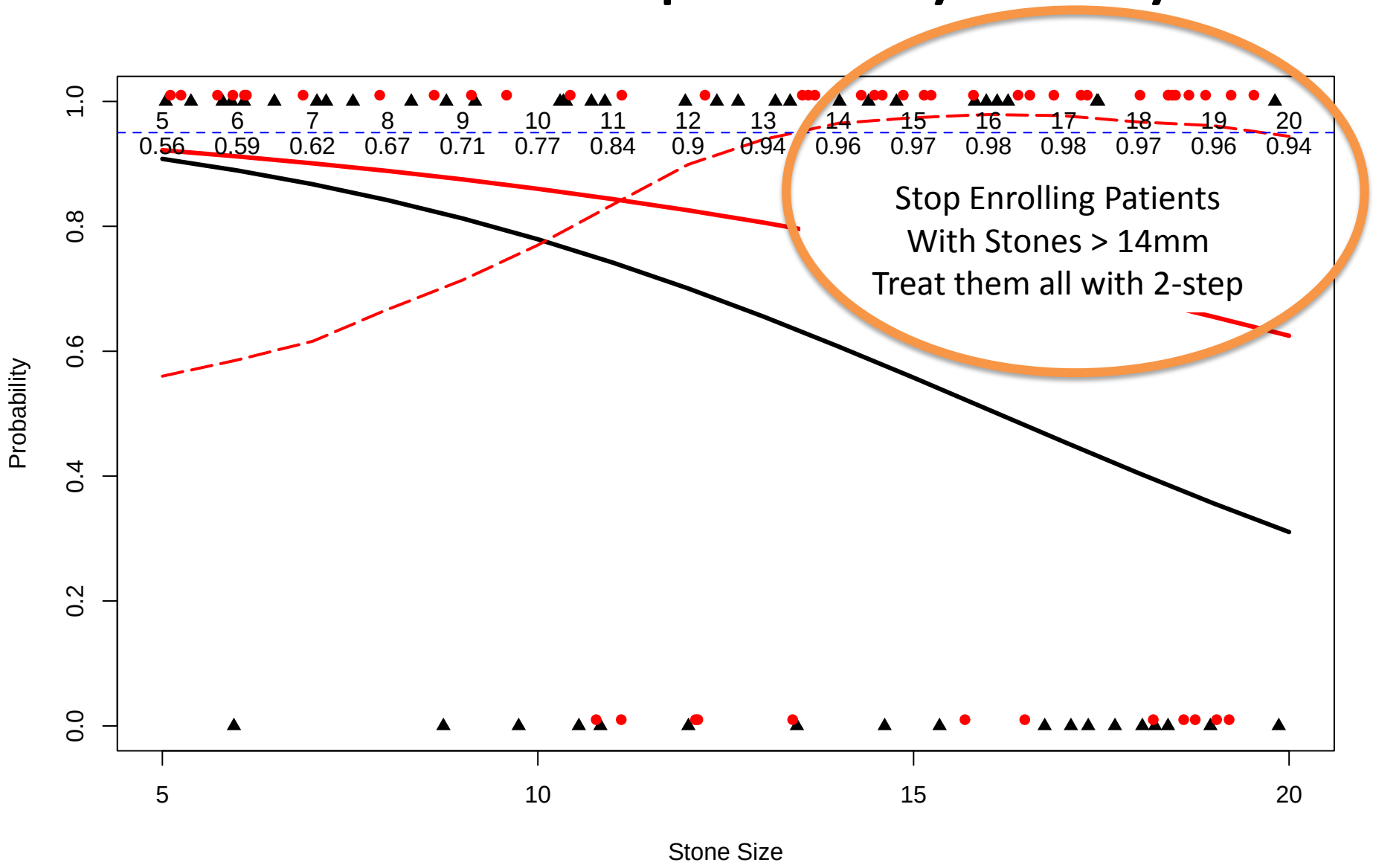
Ex1: Interim Analysis after N1=100



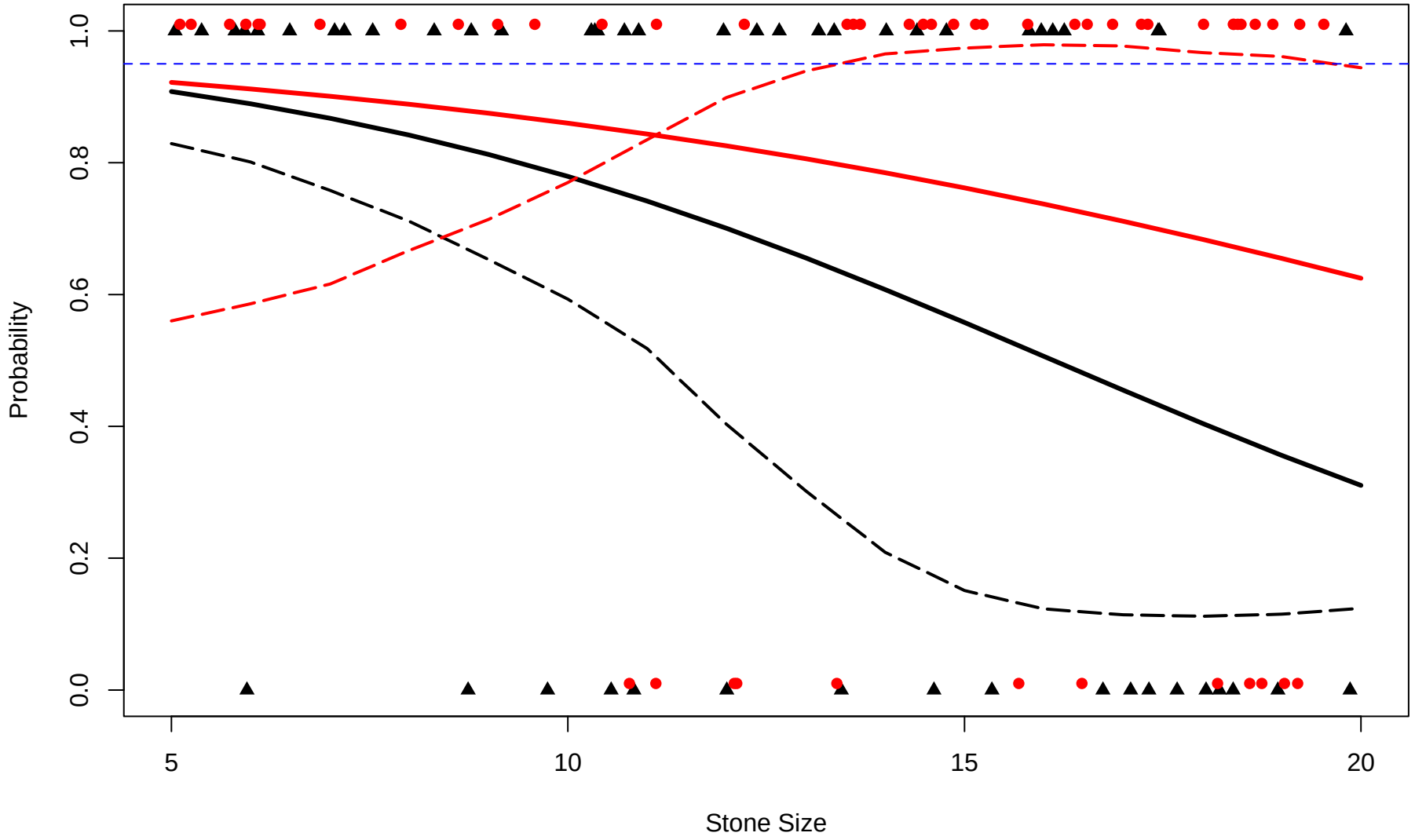
Ex1: Interim Superiority Analysis



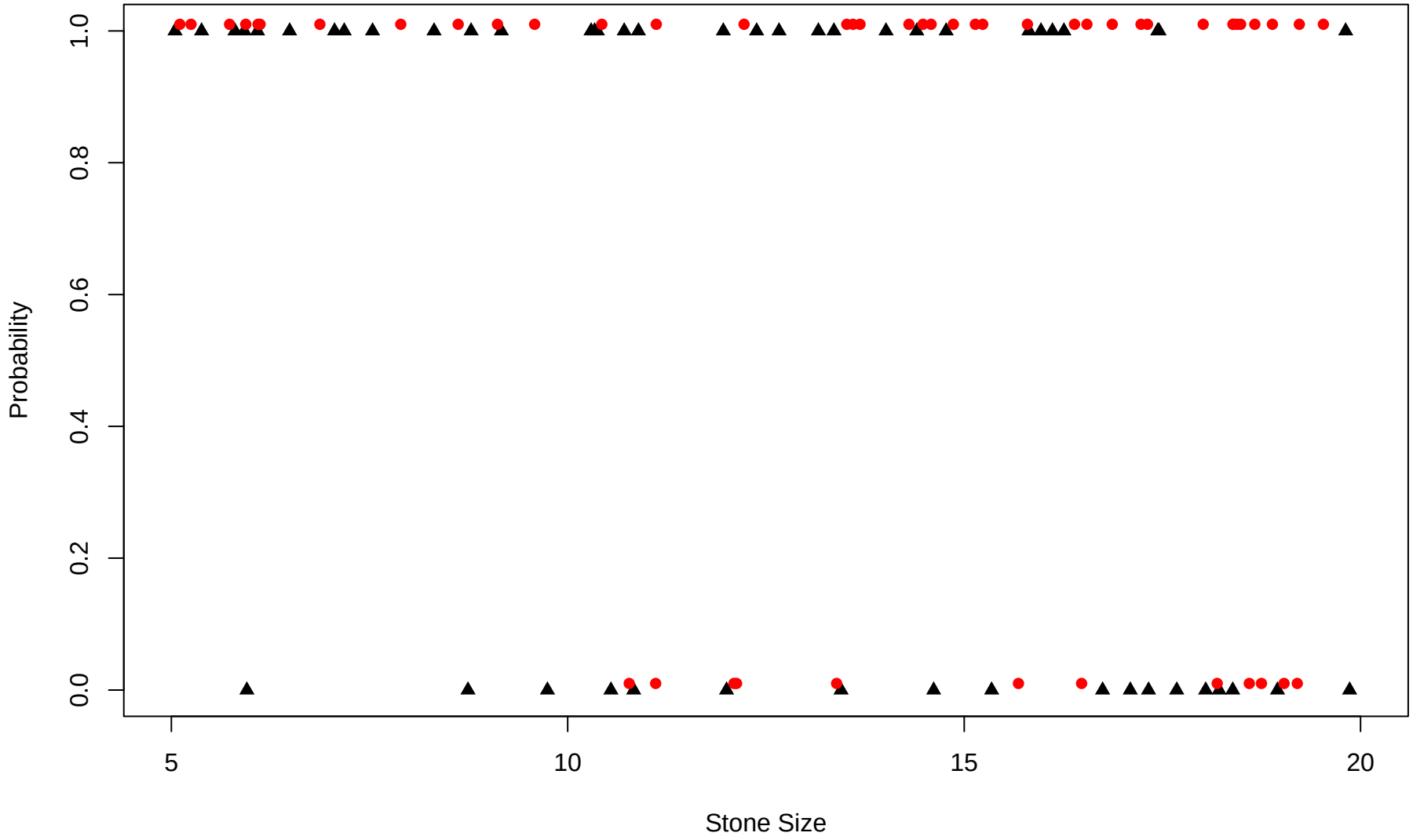
Ex1: Interim Superiority Analysis



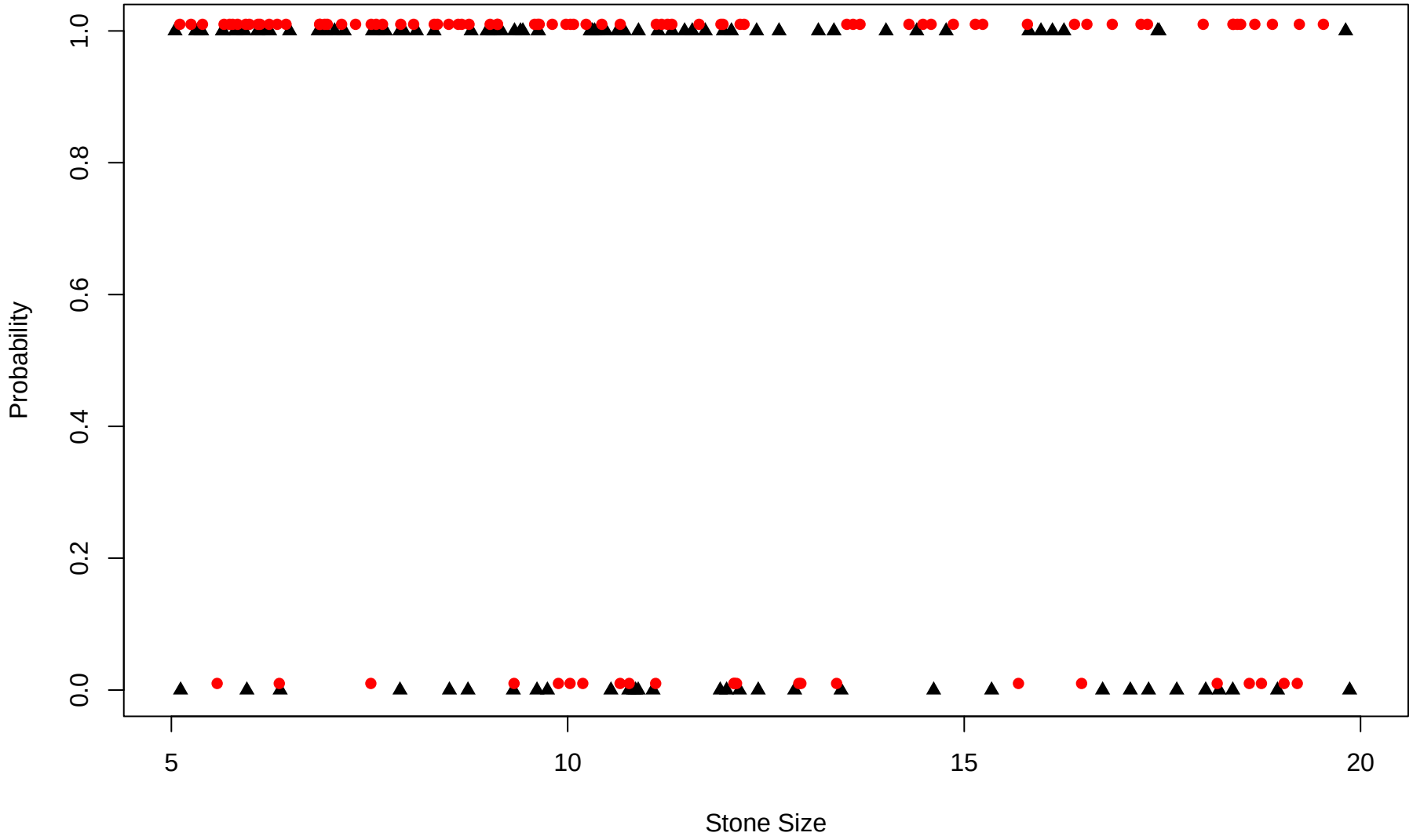
Ex1: Interim Non-inferiority Analysis



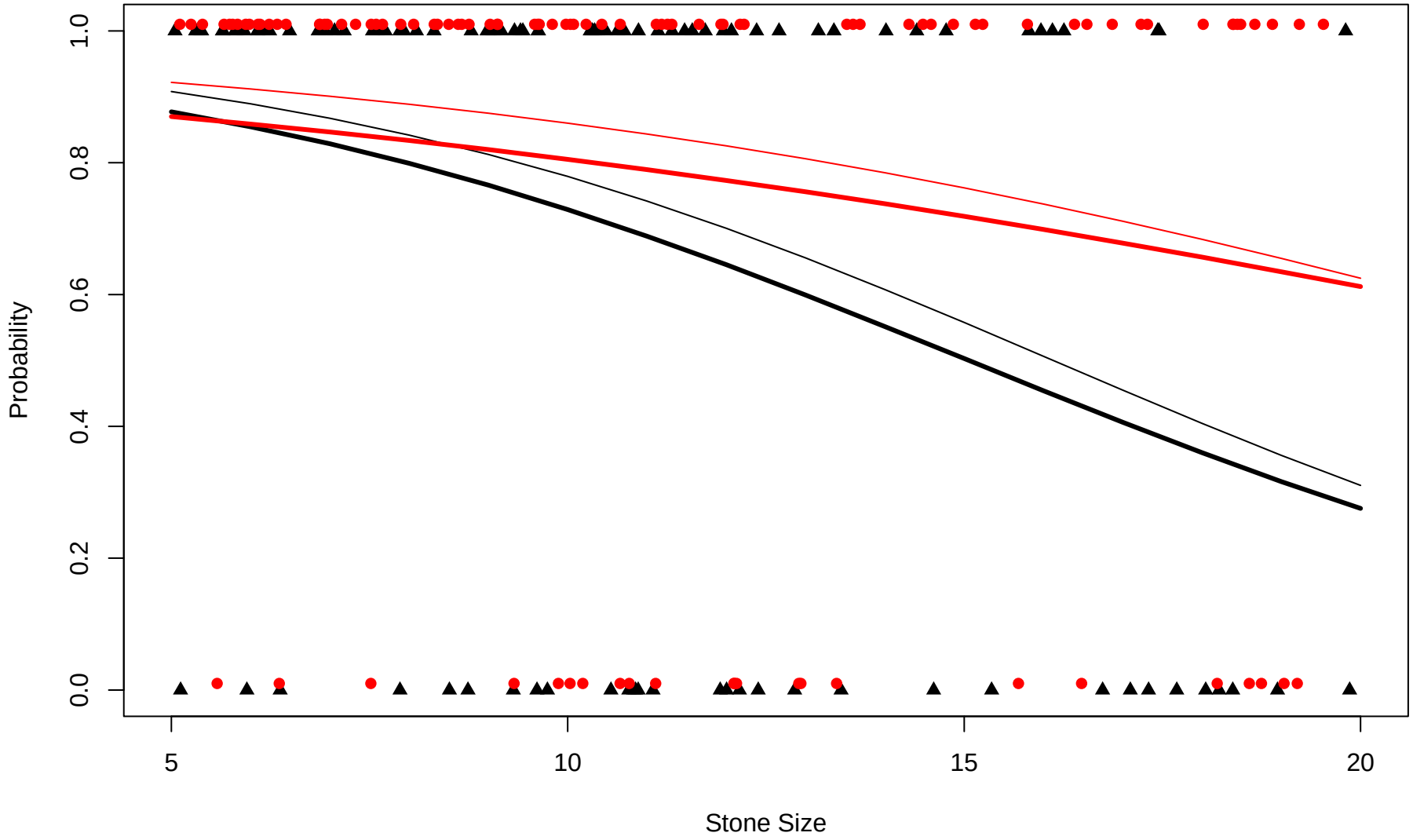
First 100 Patients
Enroll 100 more with stones 5-13



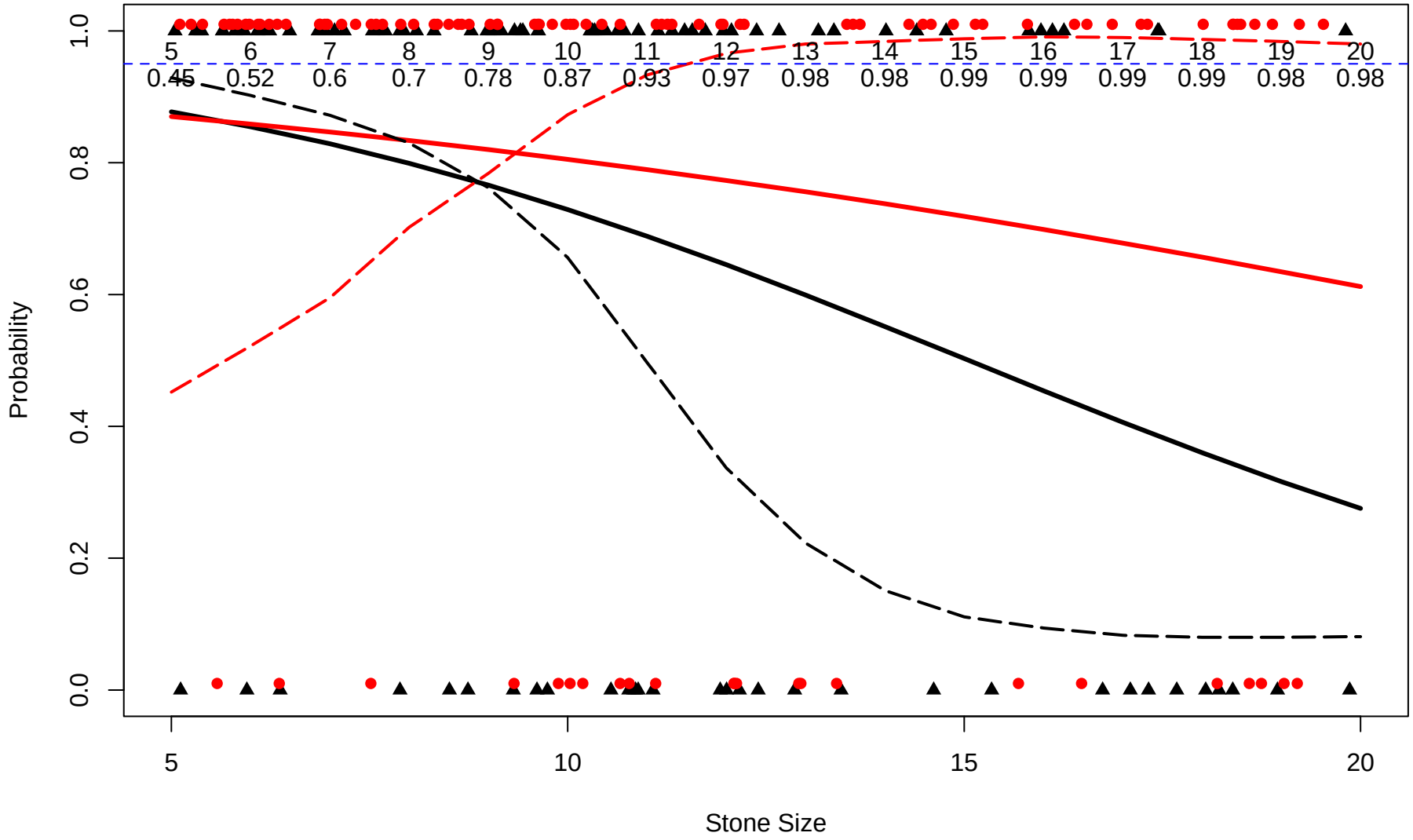
First 100 Patients
Enroll 100 more with stones 5-13



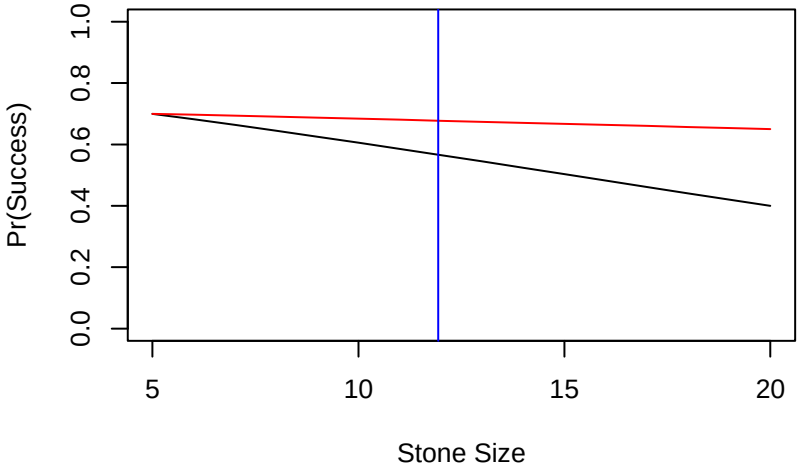
Ex1: Final Analysis after N=200

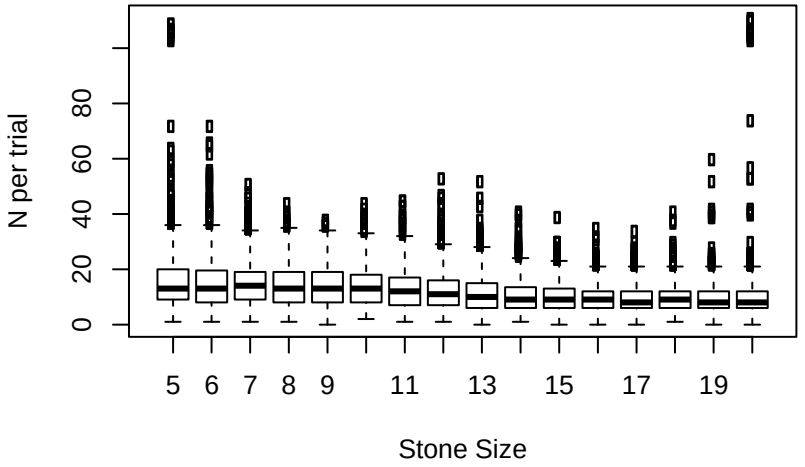


Ex1: Final Sup & NI Analysis, N=200

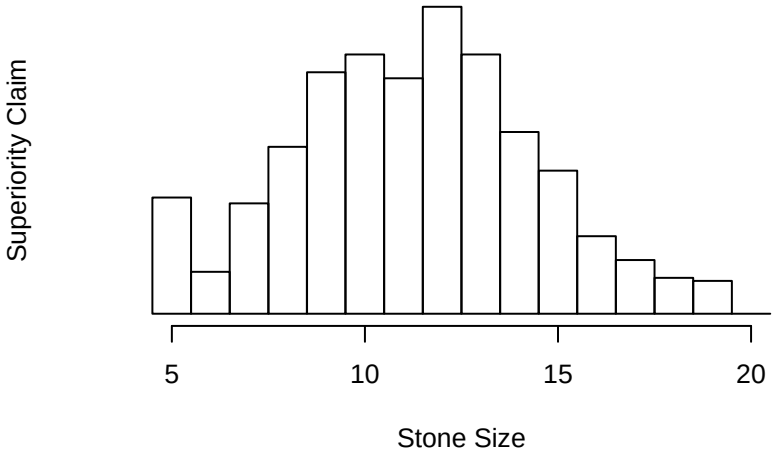


Operating Characteristics, Case 1; N=200

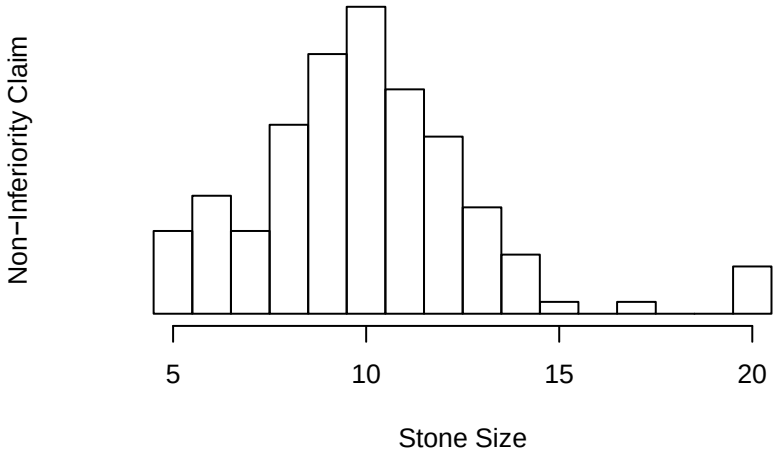




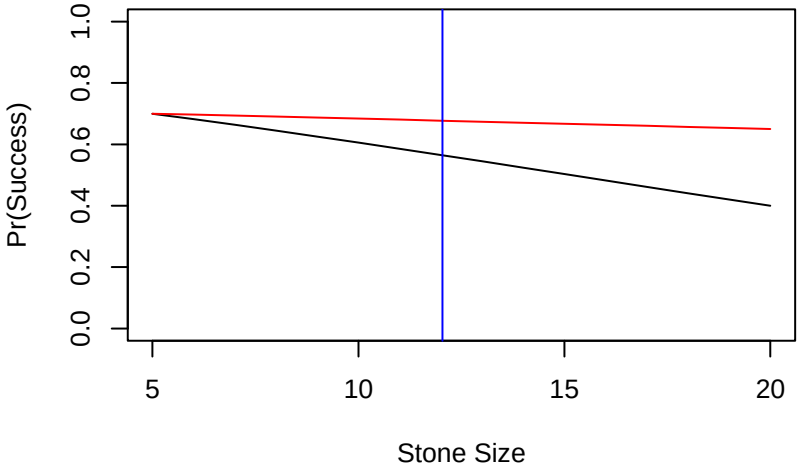
Sup Power =0.76

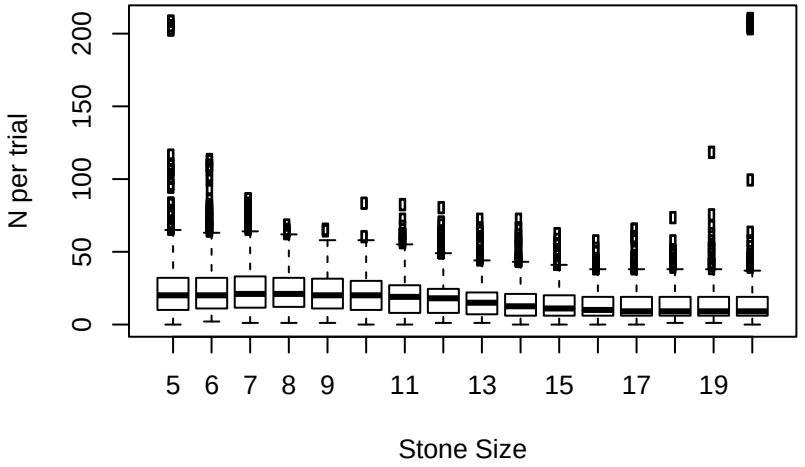


NI Power =0.14

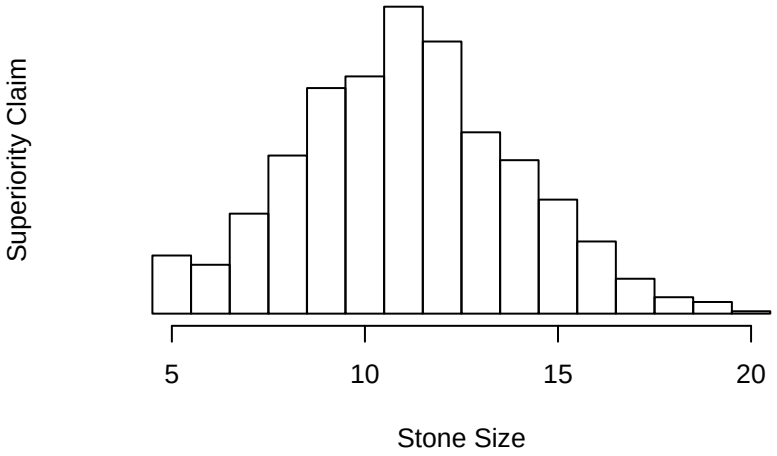


Operating Characteristics, Case 1; N=300

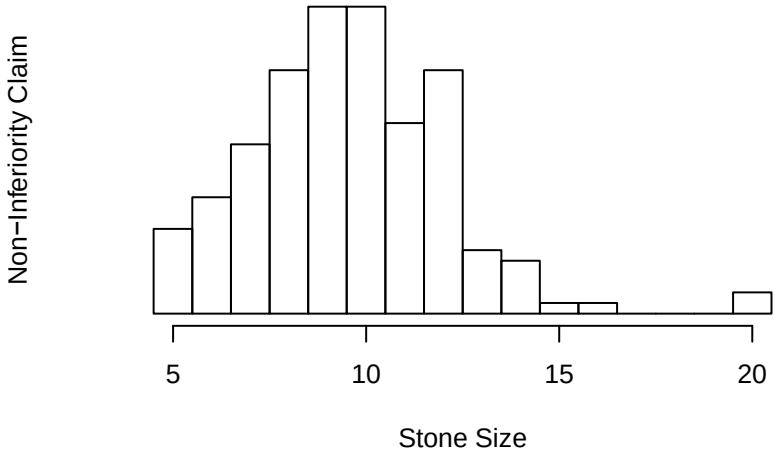




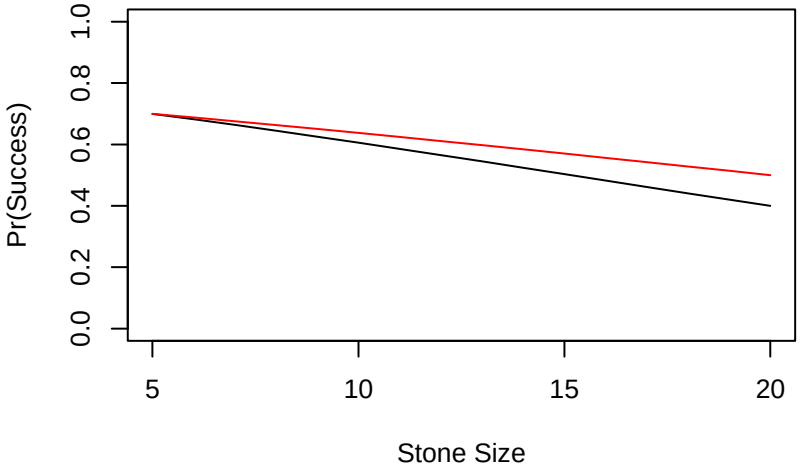
Sup Power =0.86

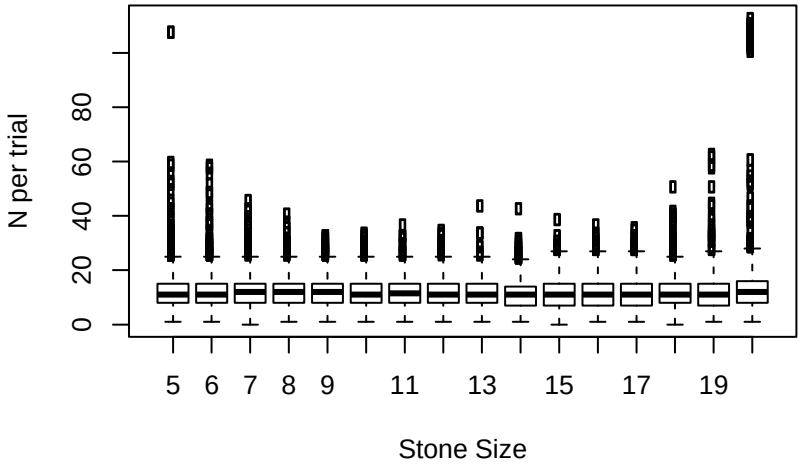


NI Power =0.17

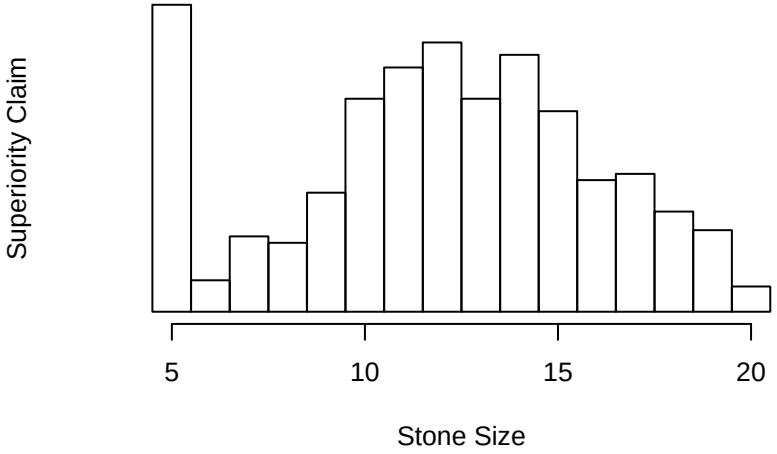


Operating Characteristics, Case 2; N=200

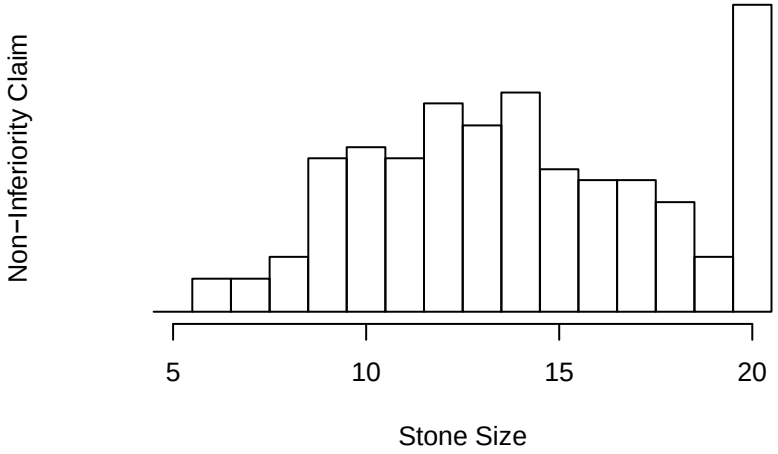




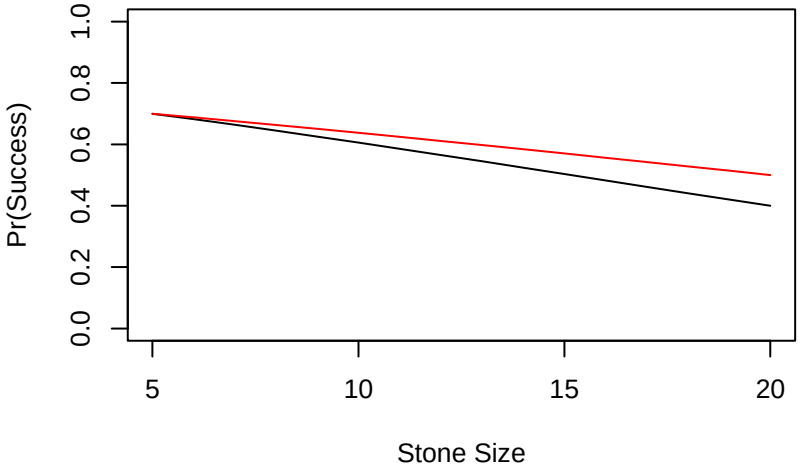
Sup Power =0.4

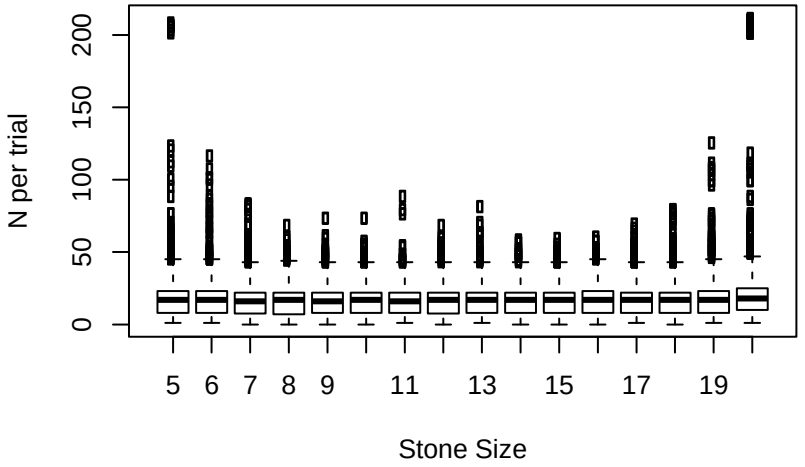


NI Power =0.19

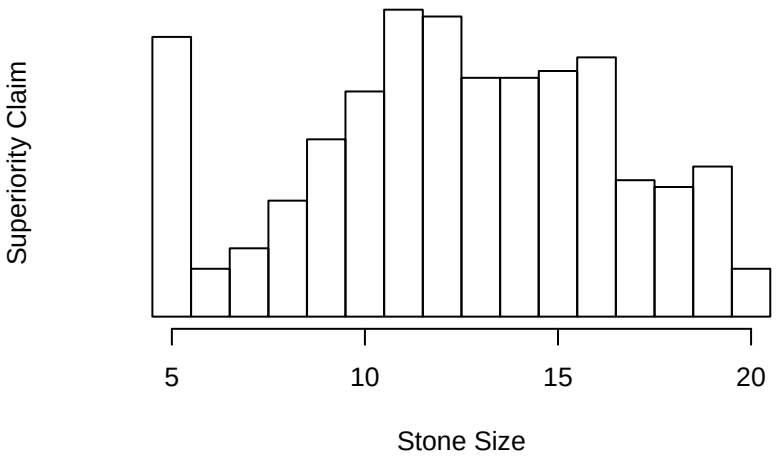


Operating Characteristics, Case 2; N=300

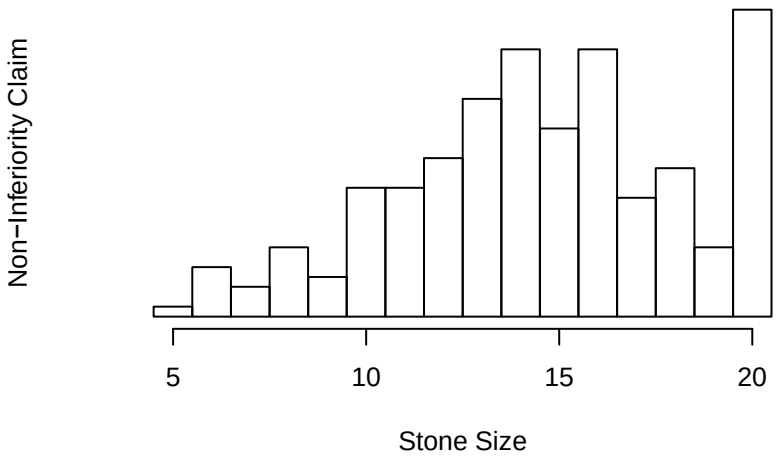




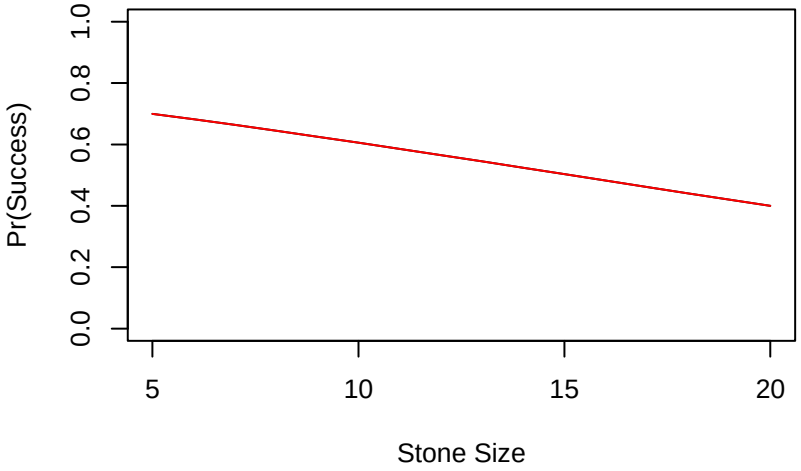
Sup Power =0.44

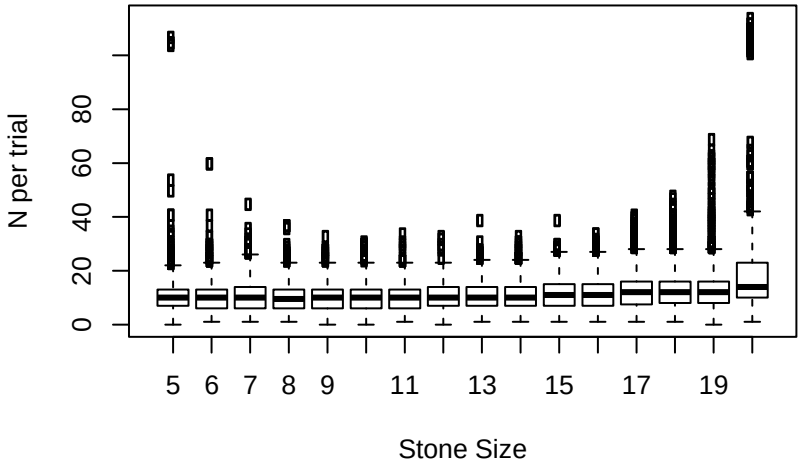


NI Power =0.22

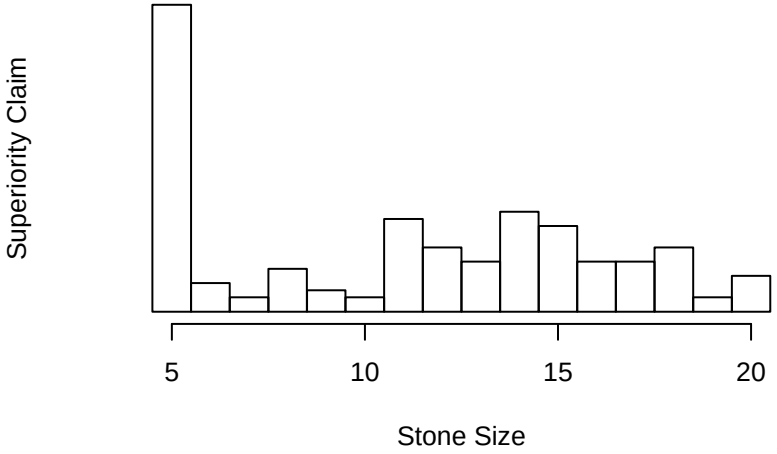


Operating Characteristics, Case 3; N=200

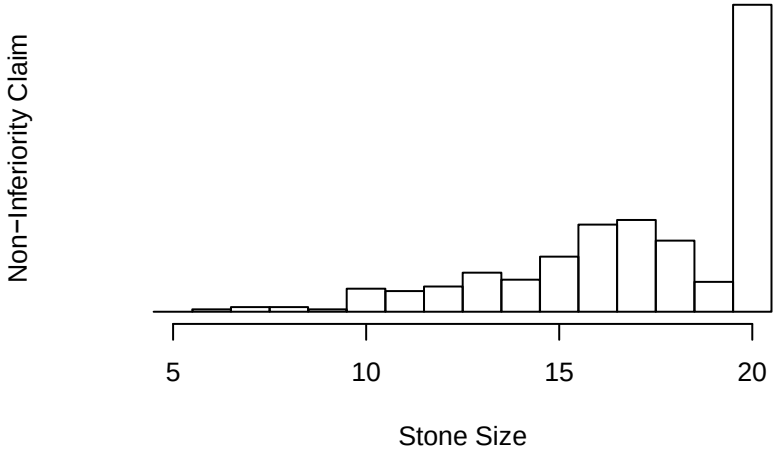




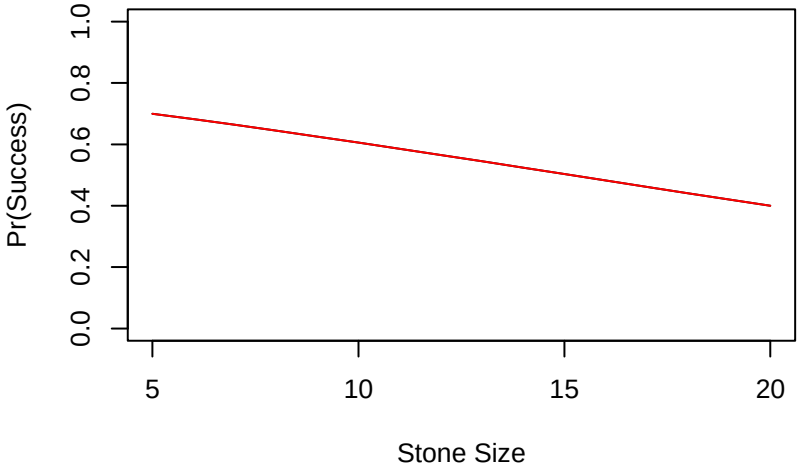
Sup Power =0.14

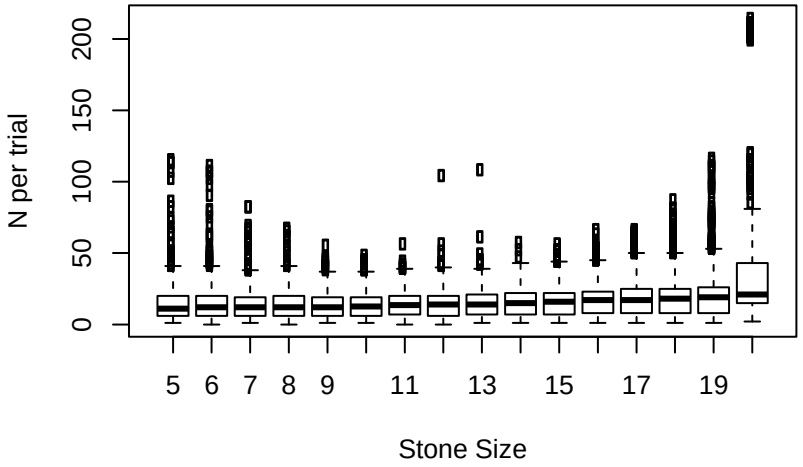


NI Power =0.35

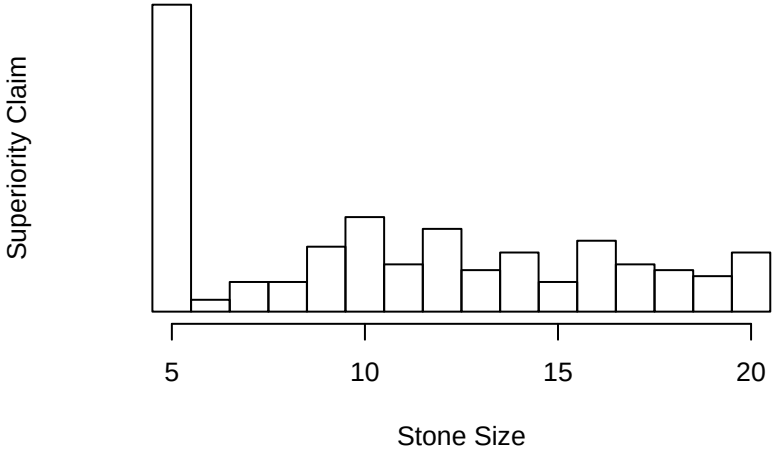


Operating Characteristics, Case 3; N=300

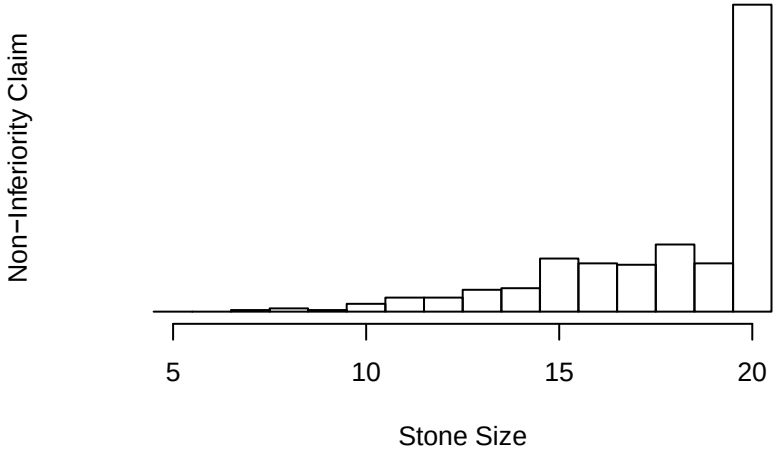




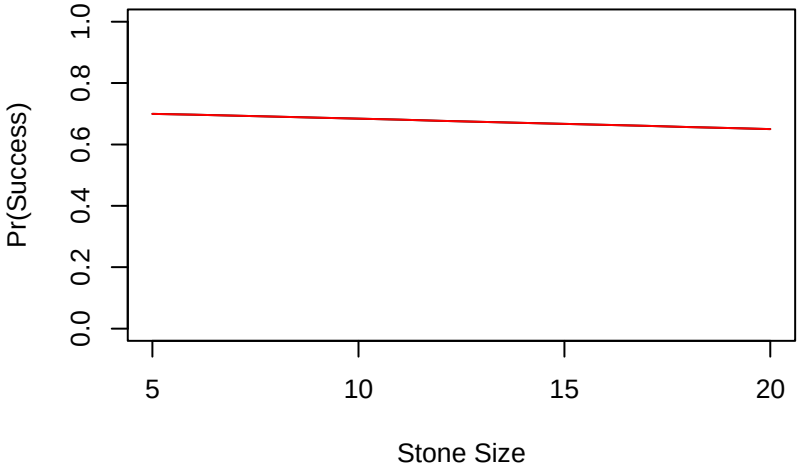
Sup Power =0.18

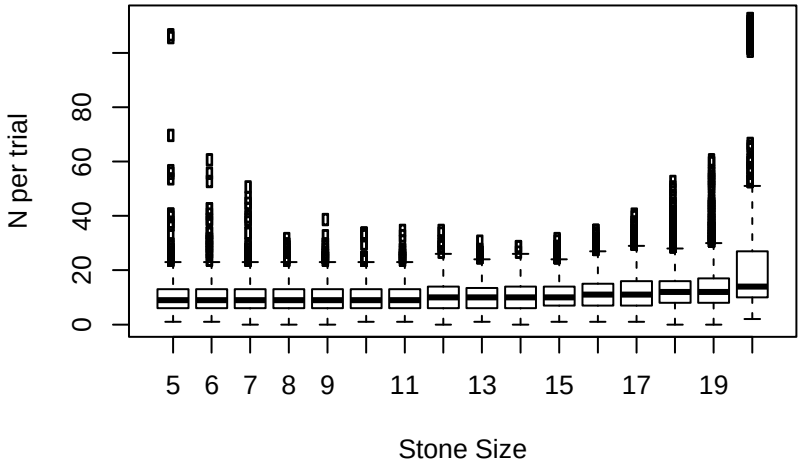


NI Power =0.42

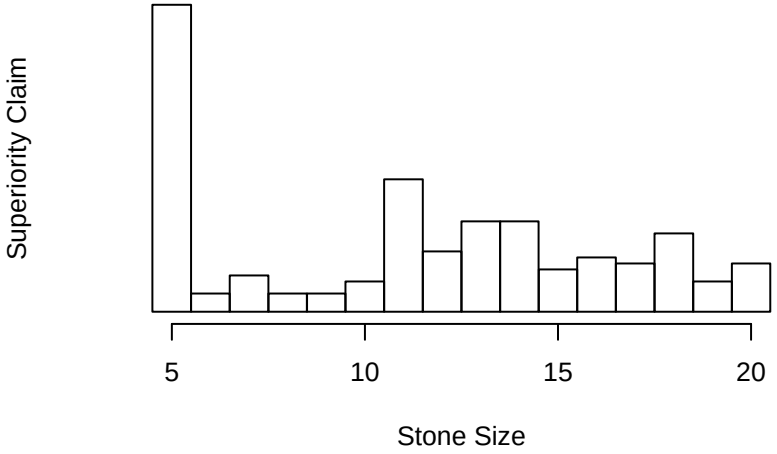


Operating Characteristics, Case 4; N=200

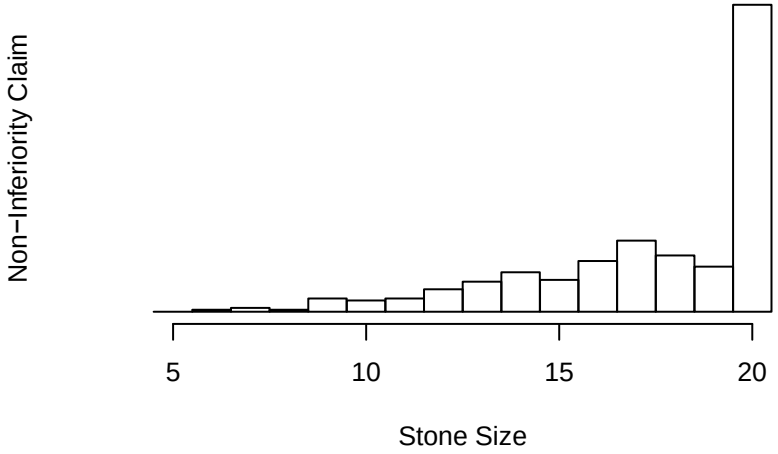




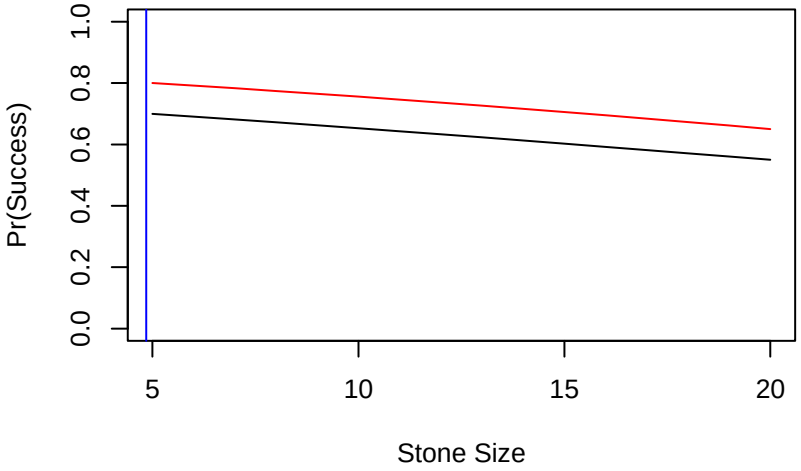
Sup Power =0.18

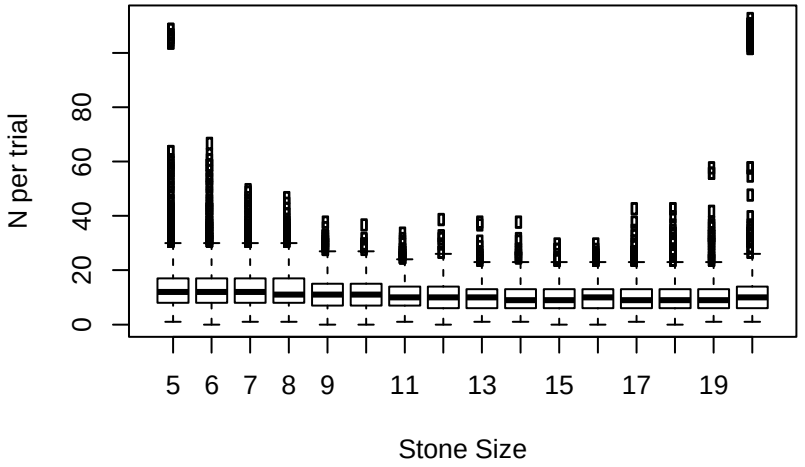


NI Power =0.37

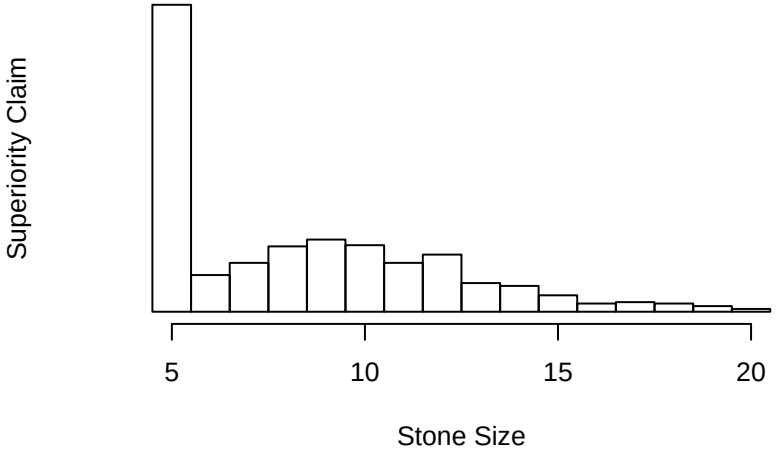


Operating Characteristics, Case 5; N=200

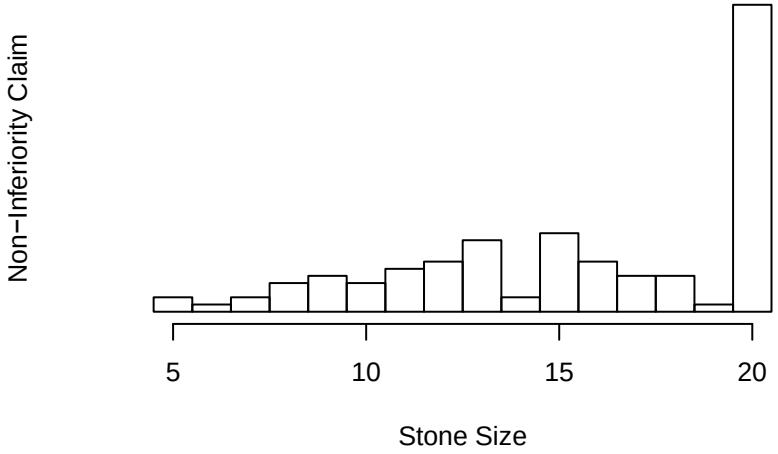




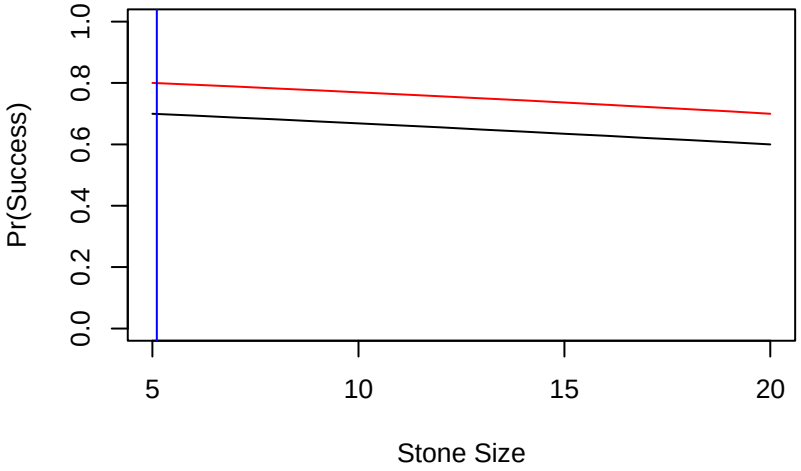
Sup Power =0.59

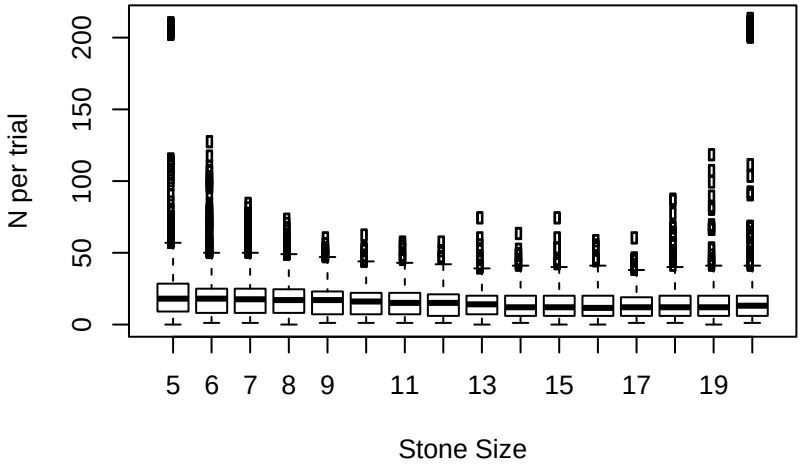


NI Power =0.12

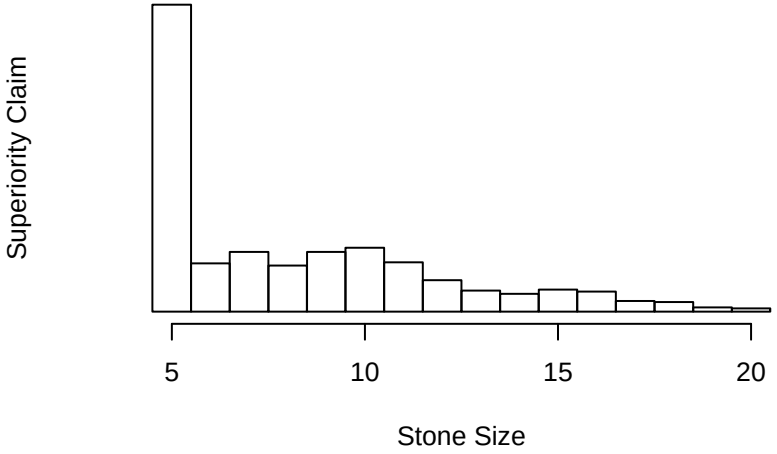


Operating Characteristics, Case 5; N=300

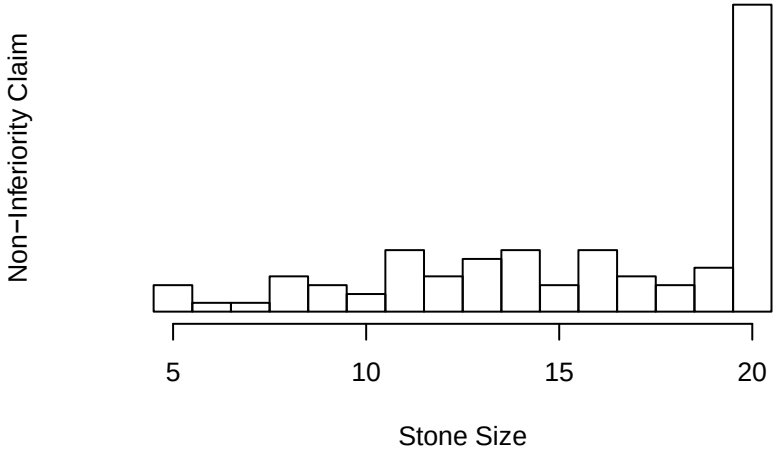




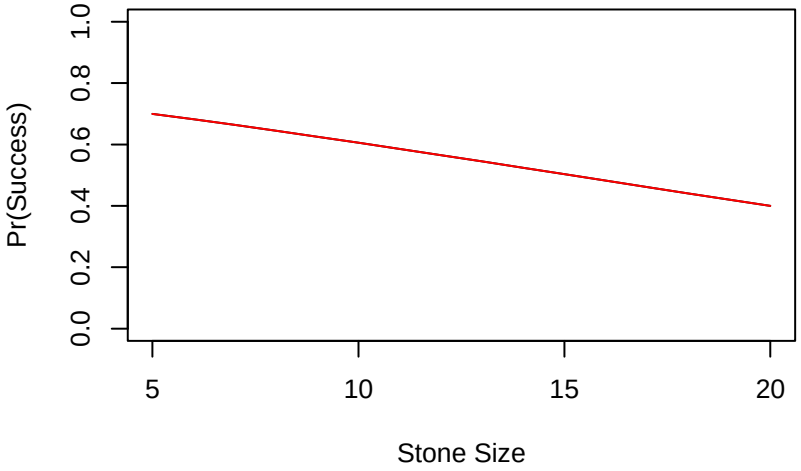
Sup Power =0.74

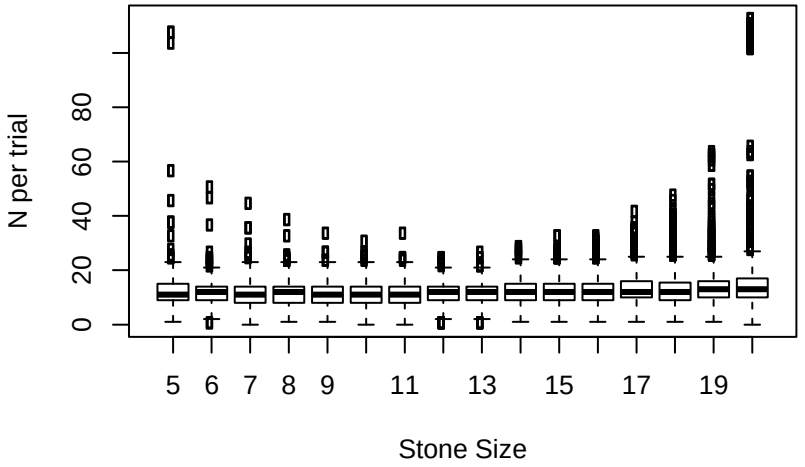


NI Power =0.1

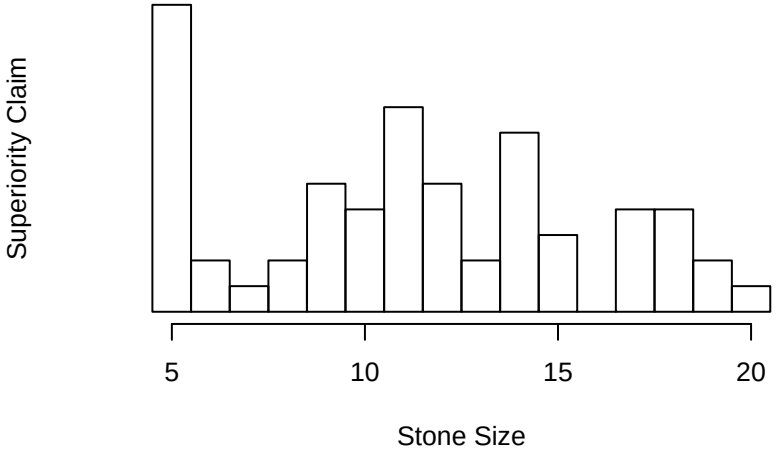


Control Type I error, .95→ .99/.98

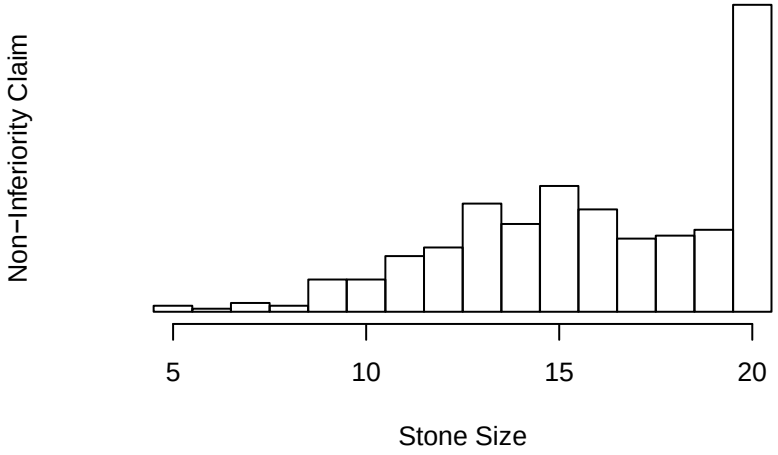




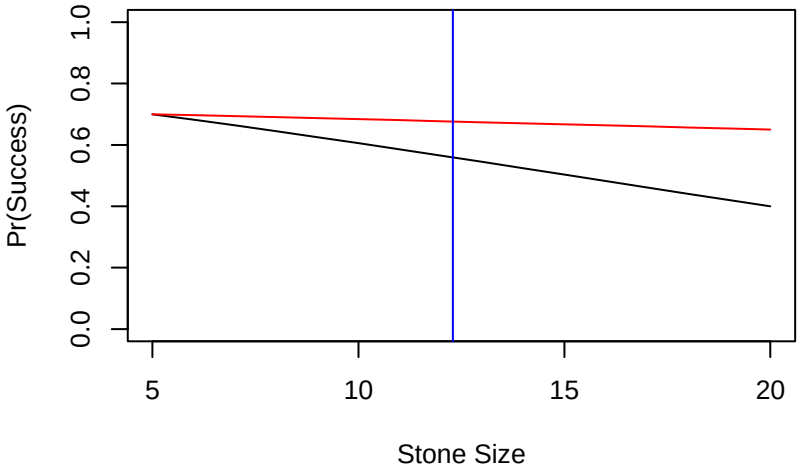
Sup Power =0.05

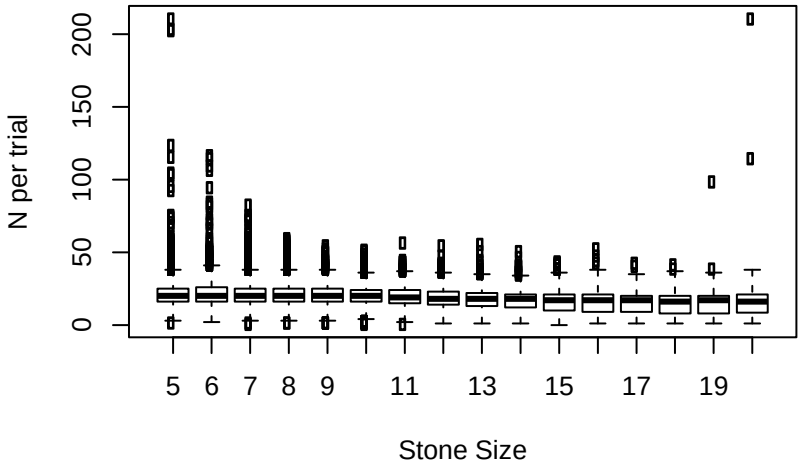


NI Power =0.4



Control Type I error, .95→ .99/.98

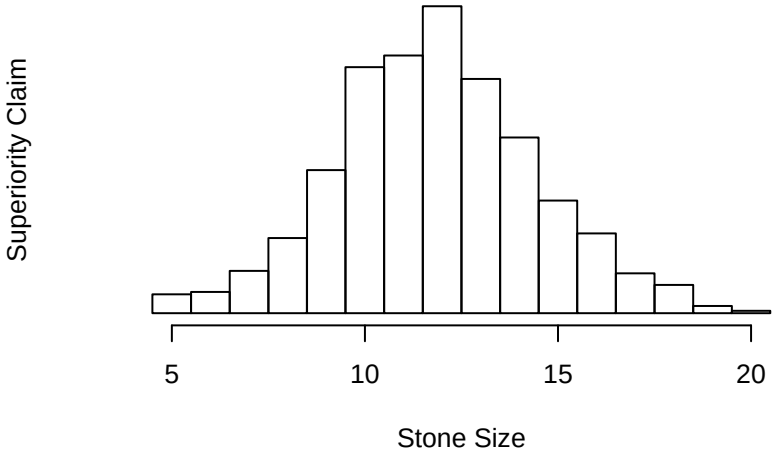


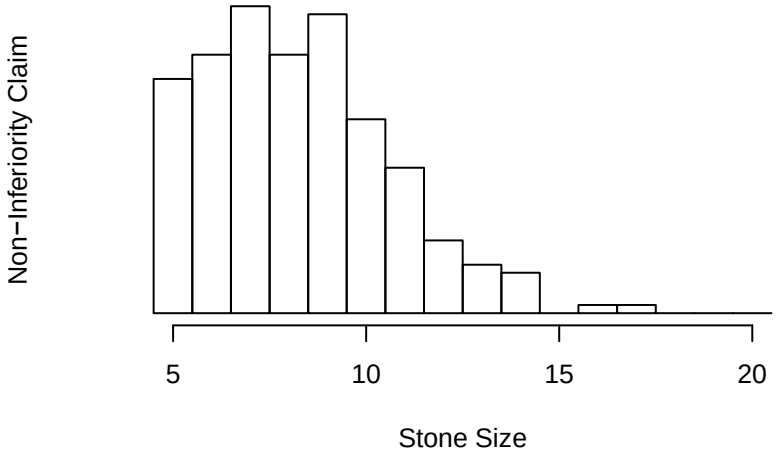


Sup Power =0.76

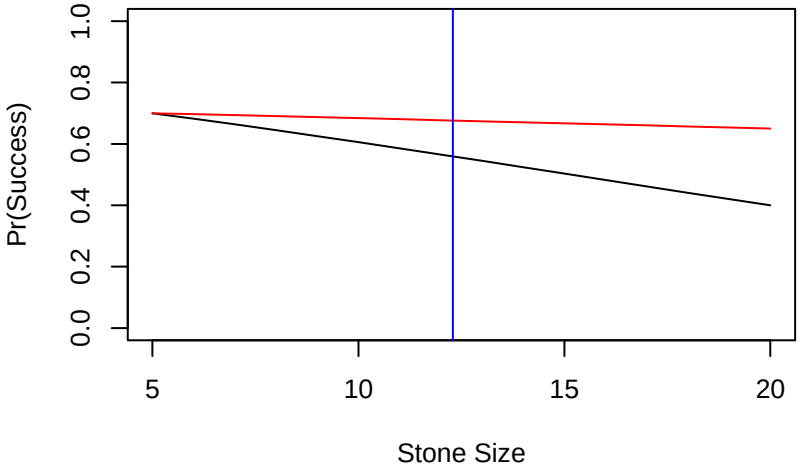
← Was 86% at 95/95

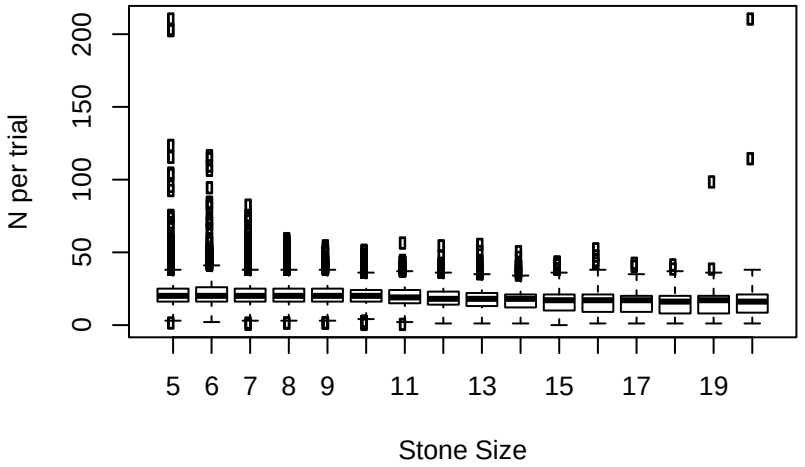
NI Power =0.23





Control Type I error, .95→ .99/.98

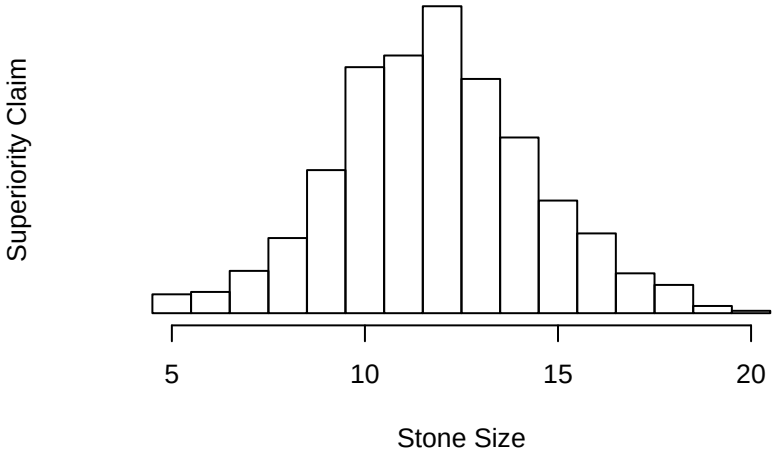


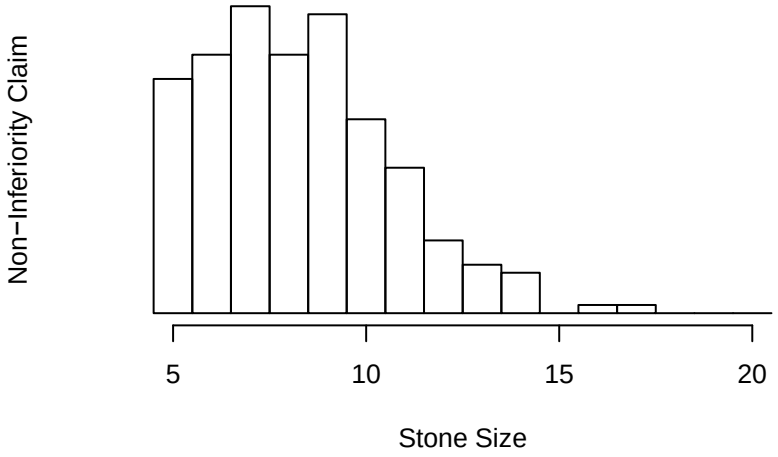


Power for n=300 Fixed Trial just comparing groups = 60%

Sup Power =0.76

NI Power =0.23





Summary

- Enrichment trial increases power & leads to personalized medicine
- Need to control Type I error via use of extensive simulation
- May need to formalize hypothesis test
  - Learn in first N1 patients
  - Choose NI region < C_1, Sup region > C_2
  - Confirm in second N2 patients
  - Hypothesis test NI in <C1 & Sup in >C2
    - Clearly state whether all patients or just second set to be used in final analysis
    - Adjust for multiplicity if the former
  - Explore choice of N1 & N2 to optimize learn vs. confirm
- Consider how operational bias may effect trial
  - Ask lots of people lots of questions