Data-Driven Operations Research Analyses in the Humanitarian Sector

Lawrence M. Wein

Graduate School of Business

Stanford University

RECENT TOPICS

PTSD in returning troops (*Mgmt Science, NY Times*)

Allocating blood for transfusions (*Transfusion*)

Space debris (*Advances in Space Research*)

Screening for childhood obesity (*Obesity, Mgmt Science*)

Ballistic imaging for solving crimes (*J Forensic Science*)

Allocating interventions to reduce childhood mortality (*PNAS*,…)

Verifying biometrics for social inclusion (*PLoS ONE*)

Reducing overcrowding in CA jails (*PLoS ONE*, *NY Times*)

Fecal transplantations (*Microbiome, PLoS ONE*)

TODAY'S TALK

Allocating food aid (randomized trial data)

Allocating food aid (nutrition program data)

Undernutrition and disease (malaria)

Optimizing health interventions (LiST model)

Assessing U.S. food assistance delivery policies

Verifying biometrics for social inclusion

A Ready-to-Use Food Allocation Policy to Reduce the Effects of Childhood Undernutrition in Developing Countries

Yan Yang, Institute for Computational & Mathematical Engineering, Stanford
Jan Van den Broeck, Centre for Int'l. Health, U. Of Bergen, Norway
Lawrence M. Wein, Graduate School of Business, Stanford University

Proceedings of the National Academy of Sciences 2013

CHILDHOOD UNDERNUTRITION

Metrics

- Weight-for-height z score (WHZ)
  - Wasting: WHZ < -2
  - Severe wasting: WHZ < -3
  - Acute (disease, food shortage) and reversible
  - 9% prevalence of wasting in Sub-Saharan Africa (< 5 years old)

CHILDHOOD UNDERNUTRITION

Metrics

- Weight-for-height z score (WHZ)
  - Wasting: WHZ < -2
  - Severe wasting: WHZ < -3
  - Acute (disease, food shortage) and reversible
  - 9% prevalence of wasting in Sub-Saharan Africa (< 5 years old)

- Height-for-age z score (HAZ)
  - Stunting: HAZ < -2
  - Chronic and hard to reverse
  - Long-term effects on adult health, education and income
  - 38% prevalence of stunting in Sub-Saharan Africa (< 5 years old)

CHILDHOOD UNDERNUTRITION

Metrics

- Weight-for-height z score (WHZ)
  - Wasting: $WHZ < -2$
  - Severe wasting: $WHZ < -3$
  - Acute (disease, food shortage) and reversible
  - 9% prevalence of wasting in Sub-Saharan Africa ($<$ 5 years old)

- Height-for-age z score (HAZ)
  - Stunting: $HAZ < -2$
  - Chronic and hard to reverse
  - Long-term effects on adult health, education and income
  - 38% prevalence of stunting in Sub-Saharan Africa ($<$ 5 years old)

Effects

- Undernutrition ($z<-2$) accounts for 3.1M deaths per year (45% of deaths among under 5's)

- Highest risk for severe ($z<-3$) cases, but more due to mild ($-2<z<-1$) and moderate ($-3<z<-2$) cases

TREATMENT FOR UNDERNUTRITION

Ready-to-Use Therapeutic Food (RUTF)

- 500 kCal/day for 3 months (e.g., Plumpy'Nut)
- Has revolutionized treatment of severe wasting (WHZ<-3)

Ready-to-Use Supplementary Food (RUSF)

- 75 - 250 kCal/day for 3 months
- Enriches existing diet
- Treats moderate or mild malnutrition, prevents severe malnutrition

TREATMENT FOR UNDERNUTRITION

RESEARCH QUESTIONS

With restricted funds (supply = 12% of demand), NGOs and governments unsure how to allocate food aid:

- Which children, based on age, sex, HAZ and WHZ?
- How much food per child?

Compare optimal policy to existing recommendations by USAID and UN’s World Food Programme:

- Depends only on age and WHZ

APPROACH

Use data set to model population-wide evolution of (HAZ,WHZ)

- 5657 children from Bwamanda (Dem. Rep. Congo) from age 6–60 months
- Bivariate ARIMA(2,1,0) structure, discrete time interval = 3 months

APPROACH

Use data set to model population-wide evolution of (HAZ,WHZ)

- 5657 children from Bwamanda (Dem. Rep. Congo) from age 6–60 months
- Bivariate ARIMA(2,1,0) structure, discrete time interval = 3 months

Use randomized clinical trial data to estimate impact of treatment (Isanaka et al., *JAMA* 2009)

- E[ΔHAZ] = 0.15, E[ΔWHZ] = 0.19 from 3 months of RUTF, lognormal

APPROACH

Use data set to model population-wide evolution of (HAZ,WHZ)

- 5657 children from Bwamanda (Dem. Rep. Congo) from age 6–60 months
- Bivariate ARIMA(2,1,0) structure, discrete time interval = 3 months

Use randomized clinical trial data to estimate impact of treatment (Isanaka et al., *JAMA* 2009)

- $E[\Delta HAZ] = 0.15$, $E[\Delta WHZ] = 0.19$ from 3 months of RUTF, lognormal

Use logistic regression to estimate death rate as a function of age, sex, HAZ and WHZ

APPROACH

Use data set to model population-wide evolution of (HAZ,WHZ)

- 5657 children from Bwamanda (Dem. Rep. Congo) from age 6–60 months
- Bivariate ARIMA(2,1,0) structure, discrete time interval = 3 months

Use randomized clinical trial data to estimate impact of treatment (Isanaka et al., *JAMA* 2009)

- $E[\Delta HAZ] = 0.15$, $E[\Delta WHZ] = 0.19$ from 3 months of RUTF, lognormal

Use logistic regression to estimate death rate as a function of age, sex, HAZ and WHZ

Optimize over proposed class of policies: Minimize DALYs lost subject to food aid budget

- 91.4% of DALYs due to childhood deaths

APPROACH

Use data set to model population-wide evolution of (HAZ,WHZ)

- 5657 children from Bwamanda (Dem. Rep. Congo) from age 6–60 months
- Bivariate ARIMA(2,1,0) structure, discrete time interval = 3 months

Use randomized clinical trial data to estimate impact of treatment (Isanaka et al., *JAMA* 2009)

- E[ΔHAZ] = 0.15, E[ΔWHZ] = 0.19 from 3 months of RUTF, lognormal

Use logistic regression to estimate death rate as a function of age, sex, HAZ and WHZ

- Death rates higher for low (HAZ,WHZ) and young age

Optimize over proposed class of policies: Minimize DALYs lost subject to food aid budget

- 91.4% of DALYs due to childhood deaths

Compare optimized proposed policy to those from WFP+USAID

BENCHMARK POLICIES

Name of Policy	Targeted		Blanket		
	WHZ	Amount	WHZ	Age	Amount
RUTF-SAM	< -3	500			
RUTF-SAM + RUSF-MAM	< -3 $\in [-3, -2)$	500 250			
RUTF-MAM	< -2	500			
RUTF-SAM + RUSF-MAM + blanket 75 kCal/day	< -3 $\in [-3, -2)$	500 250	> -2	≤ 2	75
RUTF-MAM + blanket 75 kCal/day	< -2	500	> -2	≤ 2	75
RUTF-SAM + RUSF-MAM + blanket RUSF	< -3 $\in [-3, -2)$	500 250	> -2	≤ 3	250
RUTF-MAM + blanket RUSF	< -2	500	> -2	≤ 3	250

Combination of various targeted (= function of age, WHZ) and blanket (= function of age) policies

Inspired by policies proposed in WFP (2008), USAID (2011)

PROPOSED CLASS OF POLICIES

Too hard to find solution:

- How much food to give each child based on sex, age, and (HAZ,WHZ) at months t, t-3, t-6
- State of dynamic program is 6 joint PDFs

PROPOSED CLASS OF POLICIES

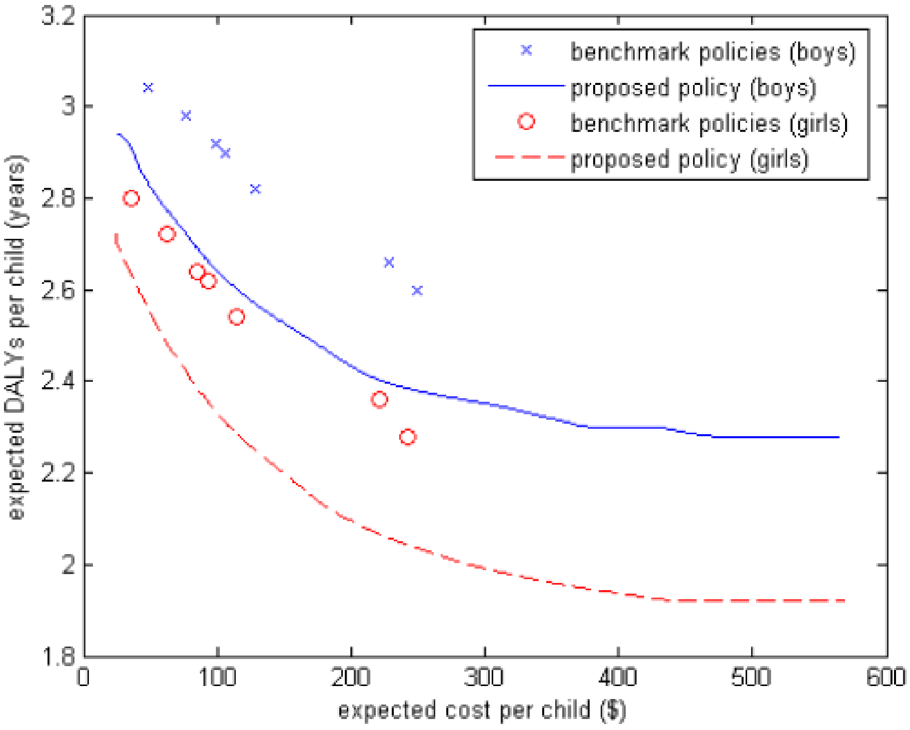
Too hard to find solution:

- How much food to give each child based on age, sex and (HAZ,WHZ) at months t, t-3, t-6
- State of dynamic program is 6 joint PDFs

Proposed class of policies:

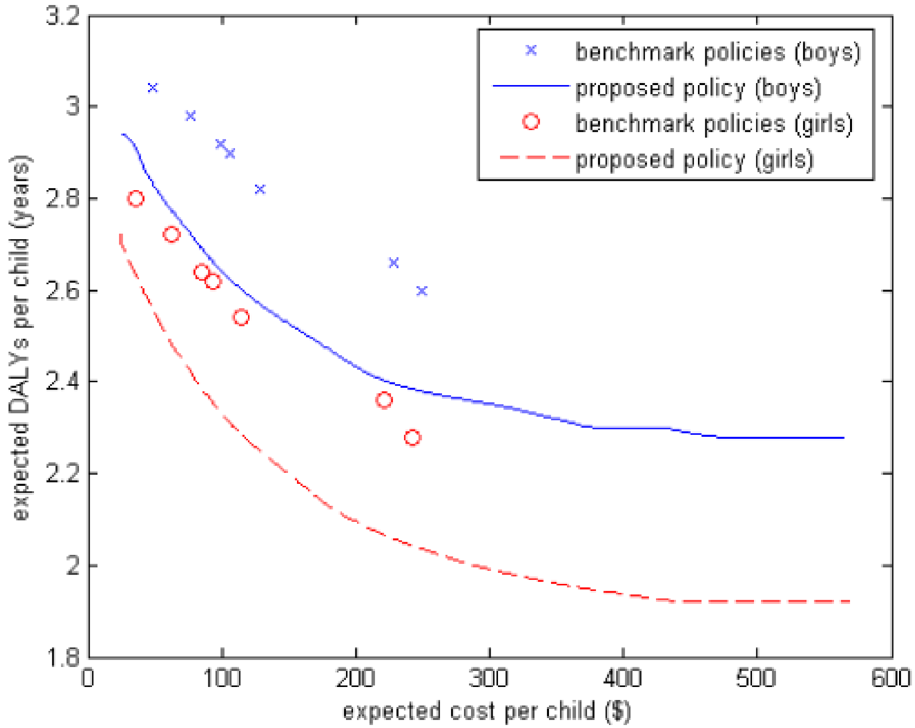
- Computable, implementable, hopefully effective
- Define child's score = contribution to objective function
  ≈ logit argument
  = Girls: 7.5HAZ + 8WHZ + age (in months)
  Boys: 7HAZ + 4WHZ + age (in months)
- Divide children into 3 buckets every 3 months:
  - Receive 500 kCal/day if score < θ_L
  - Receive x_U kCal/day if θ_L < score < θ_U
  - Receive x_L kCal/day if score > θ_U

RESULTS



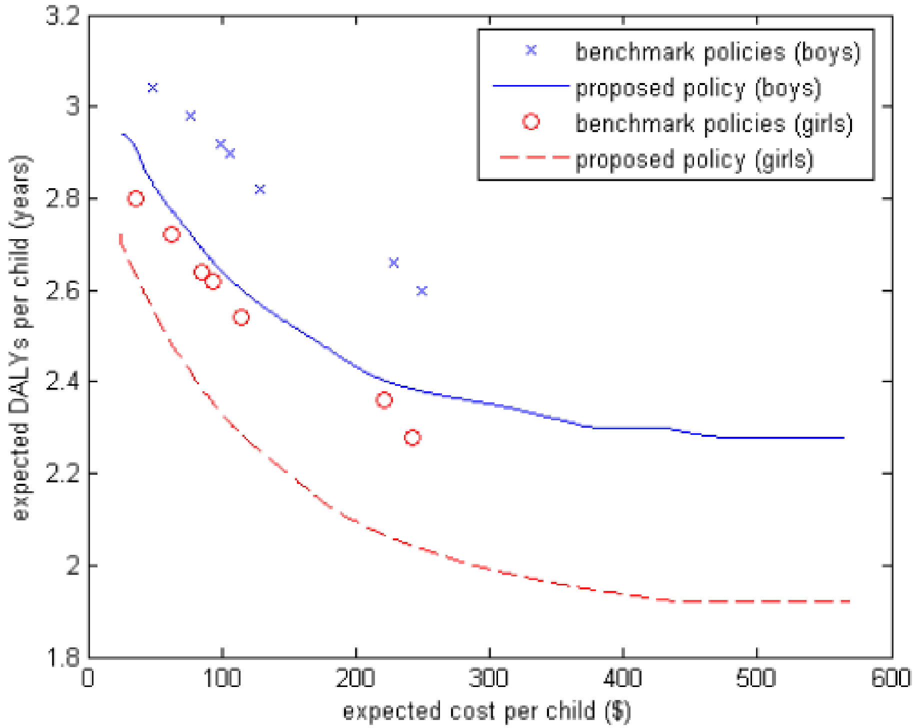
benchmark policies (boys)
proposed policy (boys)
benchmark policies (girls)
proposed policy (girls)
expected DALYs per child (years)
expected cost per child ($)
3.2
3
2.8
2.6
2.4
2.2
2
1.8
0
100
200
300
400
500
600

RESULTS



Benchmark policies can be viewed as a tradeoff curve

RESULTS



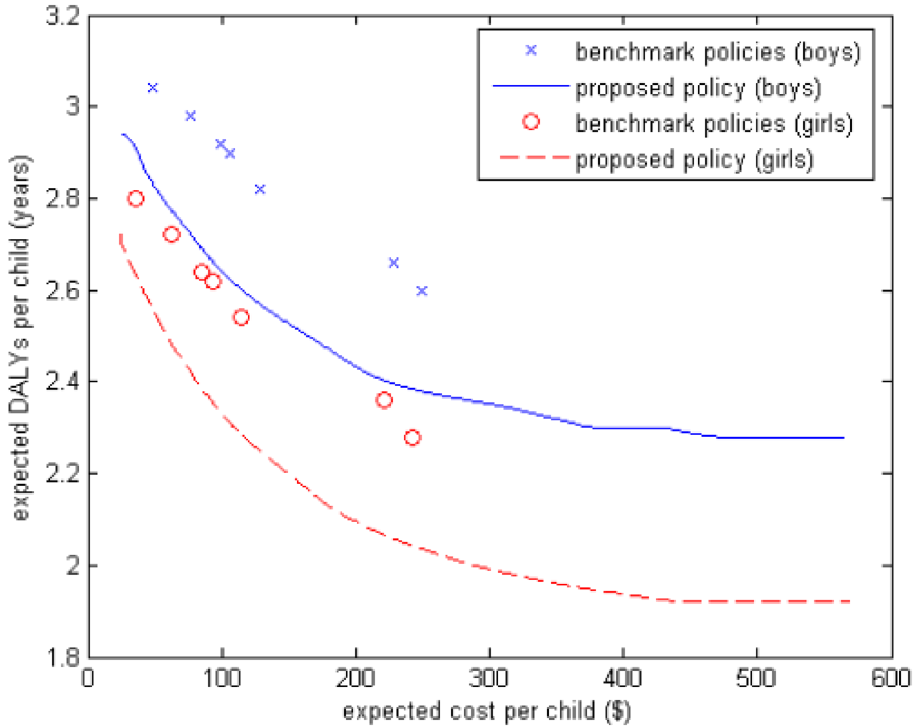
Benchmark policies can be viewed as a tradeoff curve

Proposed polices vs. benchmark policies:

9% reduction in E[DALYs] at same E[cost]

61% reduction in E[cost] at same E[DALYs]

RESULTS



Benchmark policies can be viewed as a tradeoff curve

Proposed polices vs. benchmark policies:

9% reduction in E[DALYs] at same E[cost]

61% reduction in E[cost] at same E[DALYs]

Optimal proposed policy is very simple:

500 kCal/day if score < θ

0 kCal/day if score > θ

CONCLUSION

Improvement achieved by incorporating HAZ

Regression results are consistent with odds ratios in Black (*Lancet* 2008)

Apparent contradiction: death rate for HAZ<-3 is << death rate for WHZ<-3
based on odds ratios, and yet
WHZ logistic coefficient < HAZ logistic coefficient

Explanation: 55.6% with severe wasting have severe stunting
1.8% with severe stunting have severe wasting

CONCLUSION

Improvement achieved by incorporating HAZ

Regression results are consistent with odds ratios in Black (*Lancet* 2008)

Apparent contradiction: death rate for HAZ<-3 is << death rate for WHZ<-3
based on odds ratios, and yet
WHZ logistic coefficient < HAZ logistic coefficient

Explanation: 55.6% with severe wasting have severe stunting
1.8% with severe stunting have severe wasting

Optimal “all-or-nothing” policy is at odds with blanket supplementary feeding recommendations

NEXT STEPS

Effect of treatment is weakest aspect of data

Data needs: Dependence on pre-treatment (HAZ,WHZ) and age

Probability distribution of treatment effect (interperson)

Treatment effect of supplementary food:

Is effect linear in energy consumed?

Differential targeting efficiency among food options?

Challenge: Very difficult to run randomized clinical trials

Approach: Analyze data from nonrandomized feeding programs

Paul Wise (Guatemala)

Quantifying and Exploiting the Age Dependence in the Effect of Supplementary Food for Child Undernutrition

Milinda Lakkam, Institute for Computational & Mathematical Engr, Stanford

Stefan Wager, Statistics Department, Stanford University

Paul Wise, School of Medicine, Stanford University

Lawrence M. Wein, Graduate School of Business, Stanford University

PLoS ONE 2014

GUATEMALA NUTRITION PROGRAM

Data: (height, weight, diarrhea) every 2 months from 2125 children in 17 villages

Policy: Children with weight-for-age z score (WAZ) < -2.5 get 100 kcal/day

→ regression discontinuity design

GUATEMALA NUTRITION PROGRAM

Data: (height, weight, diarrhea) every 2 months from 2125 children in 17 villages

Policy: Children with weight-for-age z score (WAZ) < -2.5 get 100 kcal/day

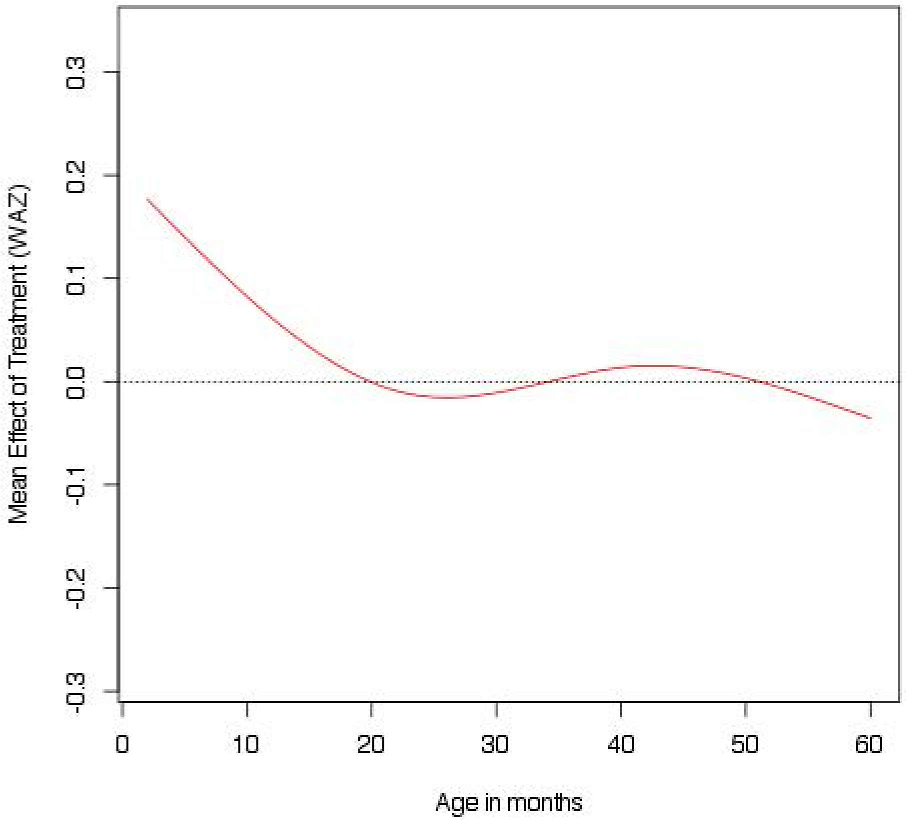
→ regression discontinuity design

Generalized Additive Model:

(E[height gain], E[weight gain], ln E[diarrhea]) = function of:

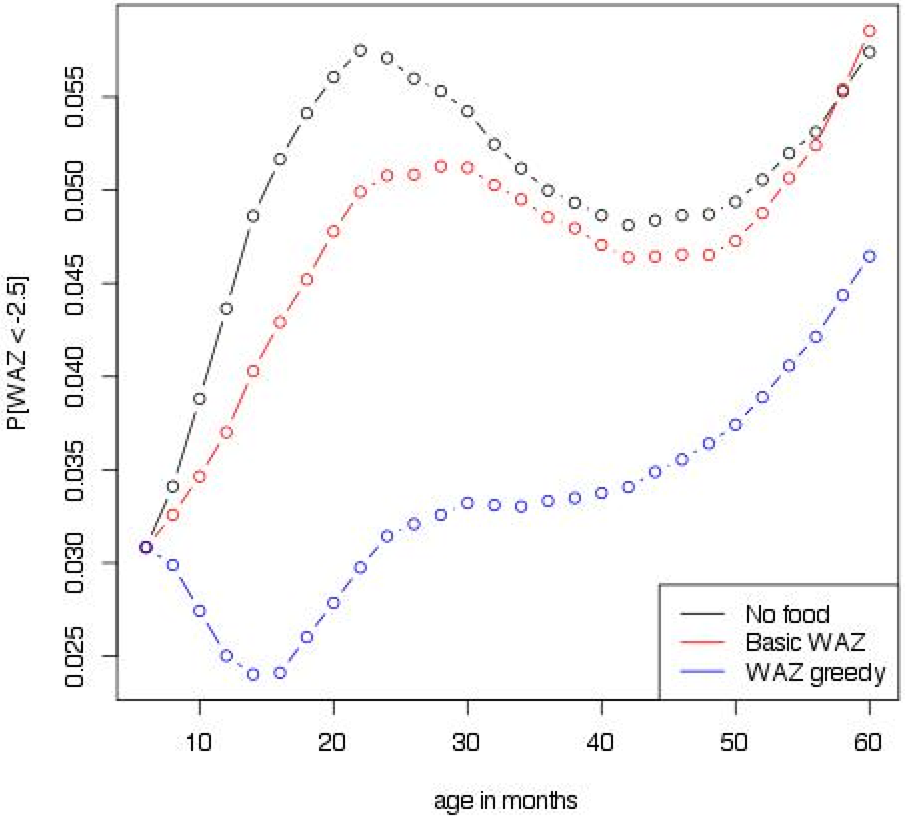
- Current (height,weight,diarrhea) global mean reversion
- Recent (height gain,weight gain,diarrhea gain) local mean reversion
- Age interaction with above terms
- Seasonality
- Treatment x age interaction

AGE DEPENDENCE IN FOOD EFFECTIVENESS



Supplementary food is only helpful during 6 – 20 months old

EXPLOITING AGE DEPENDENCE IN FOOD EFFECTIVENESS



Allocate food to minimize sum of $E[W_{t+2}]^2$ for all children with $W_t<0$

Reduces underweight severity by 14% vs current policy (which feeds 5% of children)

CONCLUSION

In low-dose (100 kCal/day), resource-constrained setting, allocating food primarily to 6-12 mo infants reduces underweight severity and stunting severity

The food allocation policy has been changed in the Guatemala nutrition program to reflect our findings

We plan to measure the change in health outcomes in an attempt to confirm our findings

Analyzing the Nutrition-Disease Nexus: The Case of Malaria

Milinda Lakkam
Institute for Computational & Mathematical Engineering, Stanford University

Lawrence M. Wein
Graduate School of Business, Stanford University

Malaria 2015

BACKGROUND

Undernutrition involved in 45% of childhood deaths

Malaria, diarrhea, pneumonia are the 3 big causes of death

Probable nutrition-disease interactions:

- Undernutrition → more susceptible to disease
- Infection → decreased nutritional status
- Undernutrition → increased mortality among infected

Possible nutrition-disease interactions:

- Undernutrition → longer infectious periods
- Infection → reduced effectiveness of supplementary food
- Undernutrition → reduced impact of infection controls

RESEARCH QUESTIONS

Can supplementary food reduce malaria morbidity and mortality?

Should Insecticide-Treated Bednets (ITNs) be targeted at children with low nutrition?

- ITNs are effective
- Demand is highly price-sensitive
- Current targeting is macro (spatial)

MODEL

Basic malaria model (Ross 1911) plus:

- Superinfection, children and adults separately

- Heterogeneous susceptibility among children
  S~Gamma(k,1/k)   (Smith, *Nature* 2005)

MODEL

Basic malaria model (Ross 1911) plus:

- Superinfection, children and adults separately

- Heterogeneous susceptibility among children

  $S \sim \text{Gamma}(k, 1/k)$ (Smith, *Nature* 2005)

  $S = U + E$

  $U \sim \text{Gamma}(k_1, 1/k)$ due to undernutrition

  $E \sim \text{Gamma}(k - k_1, 1/k)$ independent of undernutrition

  p^{th} fractile of U = $(1-p)^{th}$ fractile of WAZ

- Impact of food aid: WAZ $\rightarrow$ WAZ+ Δ (Isanaka 2009)

- Impact of ITNs (feeding cycle model of Le Menach 2007)

- Death of infected = function of WAZ (Fishman 2004)

ESTIMATING k_1

k_1 dictates proportion of heterogeneity in susceptibility that is due to undernutrition

Relative risk of WAZ < -2 for malaria =1.31(0.92,1.88) Fishman

Problem: Need to isolate impact of undernutrition on malaria from confounding factors (income, window screens, proximity to breeding areas)

More reliable estimate:

- Randomized food trial with treatment-free arm (Isanaka 2009)
- Adjusted (age,sex,seasonality,HAZ,village) odds ratio for post-treatment malaria = 0.76 (0.51,1.13)
- k_1 = 0.153 and k = 0.17 (90% of susceptibility due to undernutrition)
- Mangani (2014) has OR=0.81 (0.69,0.94) for 6-18 mo. olds

FOUR POLICIES

Targeted Food: food if WAZ < θ

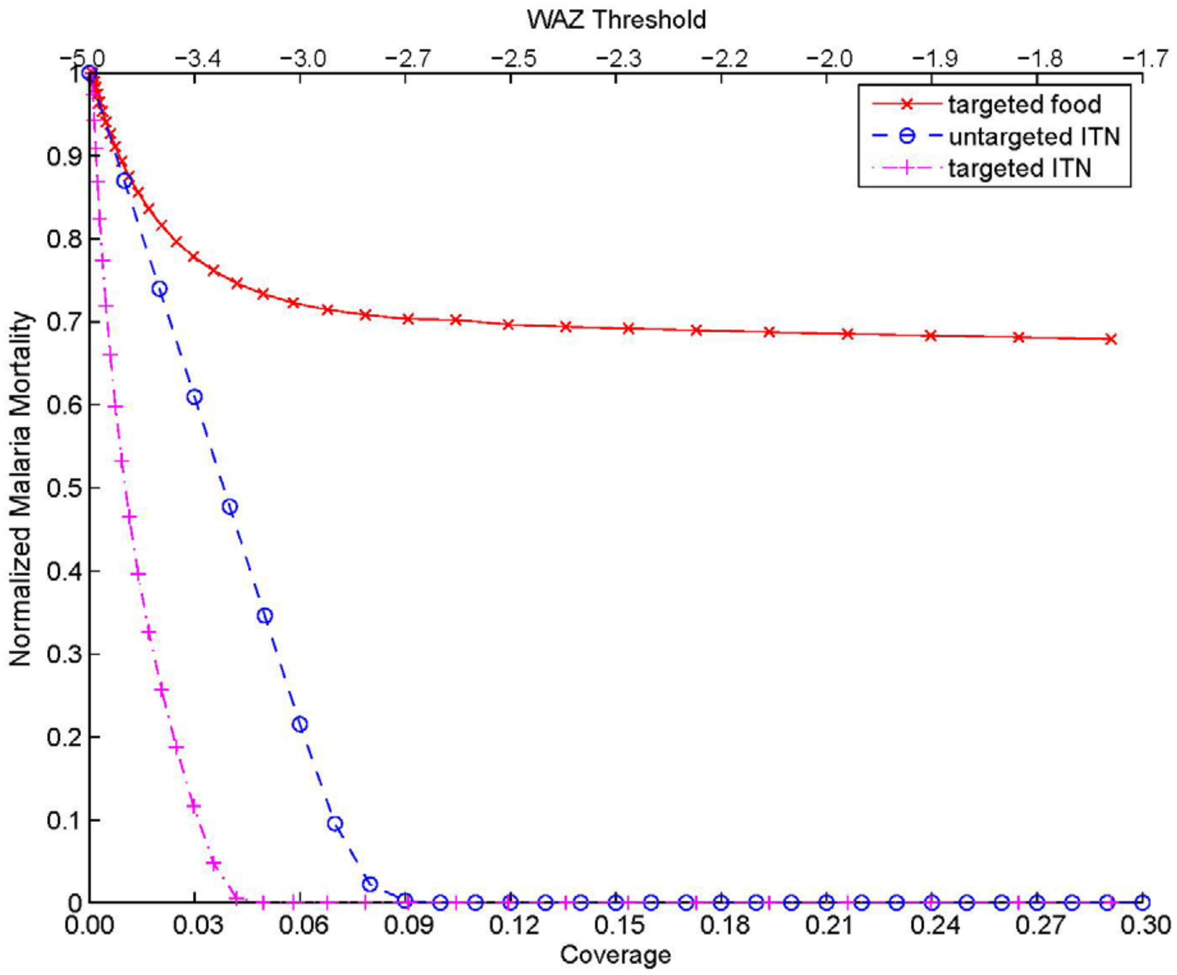
Untargeted ITN: ITN with probability p if unprotected at baseline

Targeted ITN: ITN if WAZ < θ and unprotected at baseline

Targeted food + ITN: food + ITN if WAZ < θ and unprotected at baseline

Analysis: reduction to solution to pair of equations

RESULTS: MORTALITY VS. COVERAGE



Different set of curves for combinations of:

- EIR = 1, 10, 100, 500 (measure of severity of malaria)
- Baseline ITN coverage = 20%, 50%, 80%

RESULTS

Hypoendemic (EIR = 1) setting:

- Targeting ITNs (vs. random ITNs) has most leverage
- Food helps by itself, but adds little to targeted ITN

Mesoendemic (EIR = 10) setting:

- Targeted ITN + food better than targeted food or targeted ITN

Hyperendemic (EIR > 100) setting:

- Targeted ITN policy worse than randomized ITN policy
- Food has more impact than ITNs

CONCLUSIONS

Best point estimate suggests that targeting ITNs to undernutritioned children might be effective in low-EIR setting

Results do not allow for traditional level of statistical significance, and are not robust. But:

- Identifying conflicting estimates (Fishman vs. Isanaka and Mangini) is useful
- Hypothesis is biologically plausible (undernutrition has malaria-specific immunological effects)
- Important policy implications if results are true

Possible way forward:

- Randomized cluster (i.e., at village level) ITN trial of free targeted (Z<-2) + subsidized (Z>-2) vs. subsidized (all Z) distribution in hypoendemic setting

Optimizing the Lives Saved Tool (LiST)

Apaar Sadhwani,
Management Science & Engineering Department, Stanford University

Lawrence M. Wein
Graduate School of Business, Stanford University

The Lives Saved Tool

Estimates impact on childhood mortality from a set of 77 health interventions:

- Sanitation, food, zinc, vaccines, breastfeeding, antibiotics, ITNs, micronutrients, maternal care,…

Model used widely by researchers, policy makers and program managers

Used iteratively with other tools that estimate cost and feasibility

We embed LiST into an optimization framework

- Choose intervention levels to minimize childhood mortality subject to a budget constraint
- U-shaped marginal costs: economies of scale followed by logistical or coverage problems

The Model (Winfrey et al. 2011)

i = 1,…,I interventions (I=77)
j = 1,…,J causes of death (J=25)
a = 1,…,A age bands (A=5)

$\mathbf{x_{ia}} \in [\mathbf{0}, \mathbf{1}]$ intervention level (decisions)

Problem data:
$\mathbf{p_{ja}}$ = proportion of deaths at age a due to cause j
$\mathbf{c_{ia}}$ = marginal cost rate of intervention i at age a
$\mathbf{r_{ija}}$ = mortality reduction of cause j due to intervention (i,a)
B = budget

$$\min_{x_{ia}} \quad \sum_{j=1}^{J} \sum_{a=1}^{A} p_{ja} \prod_{i=1}^{I} (1 - r_{ija} x_{ia})$$

$$\text{subject to} \quad \sum_{i=1}^{I} \sum_{a=1}^{A} c_{ia} x_{ia} \leq B$$

$$0 \leq x_{ia} \leq 1$$

Analysis I

$\mathbf{MMR_{ia}(x)}$ = marginal mortality reduction per $ of intervention i at age a when current intervention level $\mathbf{x} = (\mathbf{x_{ia}})$

$$\mathrm{MMR}_{\mathbf{ia}}(\mathbf{x}) = \frac{\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{p_{ja}}[\prod_{\mathbf{l:l\neq i}}^{\mathbf{I}}(\mathbf{1} - \mathbf{r_{lja}x_{la}})]\mathbf{r_{ija}}}{\mathbf{c_{ia}}}.$$

$\mathbf{MMR_{ia}(x)}$ (i) is constant as $\mathbf{x_{ia}}$ increases if other age a interventions are fixed

(ii) decreases as other interventions for age a are increased

Heuristic algorithm achieves local optima:

(i) Invest in unsaturated $\mathbf{x_{ia}}$ with highest $\mathbf{MMR_{ia}(x)}$ until $\mathbf{x_{ia}}$ = 1, or budget B is reached

(ii) Swap interventions if another $\mathbf{MMR_{ia}(x)}$ decreases to below Lagrange multiplier λ

Analysis II

$\mathbf{MMR_{ia}(x)}$ = marginal mortality reduction per $ of intervention i at age a at current intervention level $\mathbf{x = (x_{ia})}$

$$\mathrm{MMR}_{\mathbf{ia}}(\mathbf{x}) = \frac{\sum_{\mathbf{j=1}}^{\mathbf{J}} \mathbf{p_{ja}}[\prod_{\mathbf{l:l\neq i}}^{\mathbf{I}}(\mathbf{1 - r_{lja}x_{la}})]\mathbf{r_{ija}}}{\mathbf{c_{ia}}}.$$

Simplex-type algorithm achieves global optima:

- All vertices that are not globally optimal have a neighboring vertex that has lower mortality

Constant or nonincreasing marginal costs → at most one fractional solution

With U-shaped marginal costs, many interventions can be fractional

Comments

Next step is to apply in setting with real data

LiST is currently viewed as a black box

requires user to propose mix of interventions

Goals:

- Structural insights

  $\mathbf{MMR_{ia}}(\mathbf{x})$ quantifies cost-effectiveness of intervention

  Interventions that are too similar will cut into each other's cost-effectiveness
- Help LiST users identify better mix of interventions in static and dynamic settings

Estimating the Impact of U.S. Food Assistance Delivery Policies on Child Mortality in Sub-Saharan Africa

Alex Nikulkov and Lawrence M. Wein

Graduate School of Business, Stanford University

Christopher B. Barrett, Dyson School, Cornell University

Andrew Mude, Intl. Livestock Research Institute, Kenya

PLoS ONE 2016

Background

Every country except the U.S. donates food assistance via cash-based interventions: cash transfers, food vouchers, and Local and Regional Procurement (LRP)

U.S. = 30% is cash-based

70% transoceanic shipments of food:

- Longer lead times
- More expensive (50% U.S. carriers)

Mude, Barrett et al. (2009)

- Forecast proportion of undernutritioned children (MUAC-Z) in Northern Kenya in non-emergency setting in t+1 and t+3 mo:
- Variables = past proportions, herd sizes, herd mortality, food aid, rain, vegetation index, forage, month
- Combination of satellite and survey data

Research Question

How much does childhood mortality decrease as cash-based proportion increases from 65% (currently) to 100% (U.S. changes policy)?

Approach

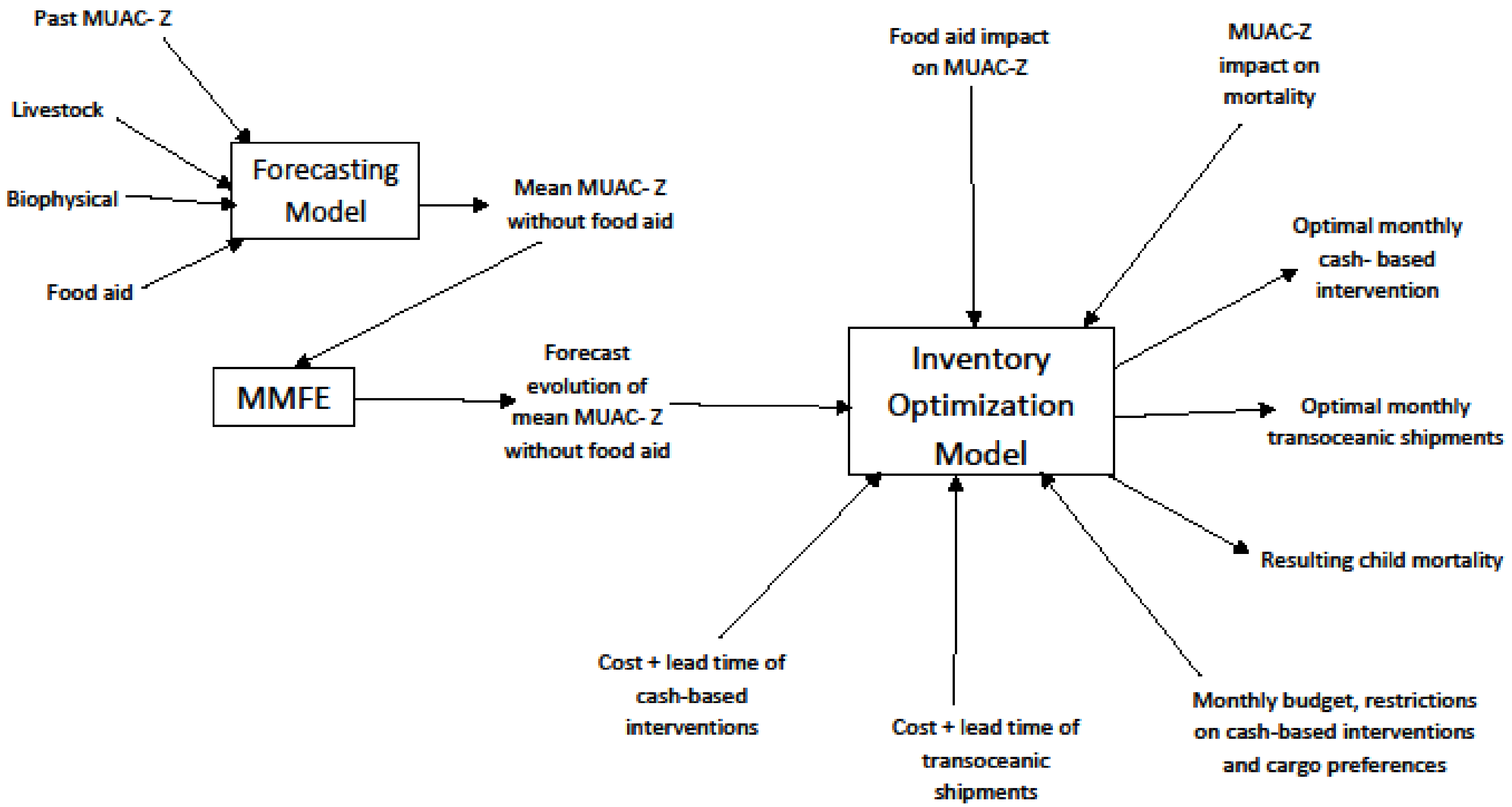
Model the accuracy of the Mude-Barnett forecasting tool using Martingale Model of Forecast Evolution (MMFE)

- Add seasonality
- Allow correlated forecast updates (for given end date)

Embed in dynamic program for monthly shipments:

- Two modes of shipment: cash (fast) and non-cash (slow)
- Constraint on minimum amount of cash-based assistance
- State = (remaining budgets, pipeline of orders, forecasts)
- Minimize childhood mortality
  - Food increases Z values
  - Z values map into mortality
- Solve using approximate dynamic programming (inspired by closed-form solution to single-mode version of problem)

Approach



Past MUAC- Z
Livestock
Biophysical
Food aid
Forecasting Model
Mean MUAC- Z without food aid
MMFE
Forecast evolution of mean MUAC- Z without food aid
Food aid impact on MUAC-Z
MUAC-Z impact on mortality
Inventory Optimization Model
Optimal monthly cash- based intervention
Optimal monthly transoceanic shipments
Resulting child mortality
Cost + lead time of cash-based interventions
Cost + lead time of transoceanic shipments
Monthly budget, restrictions on cash-based interventions and cargo preferences

Results

Child mortality decreases by 16.2% if U.S. switches to cash-based assistance

Child mortality decreases by 1.1% if U.S. only eliminates U.S.-flag vessel restriction

Focus should be agriculture lobby, not maritime lobby

Great majority of improvement from cash-based approach is due to reduced cost, not to reduced delivery lead times

Analyzing Personalized Policies for Online Biometric Verification

Apaar Sadhwani, Management Science & Engineering Dept, Stanford U.
Yan Yang, Institute for Computational & Mathematical Engr, Stanford U.
Lawrence M. Wein, Graduate School of Business, Stanford University

PLoS ONE 2014

Unique Identification Authority of India (UIDAI)

Goal: Enroll all 1.2 billion residents to improve social inclusion (e.g., health care, government services)

Capture 10 fingerprint images and 2 iris images at enrollment

UIDAI has carried out extensive verification (1-to-1 matching) experiments

Ramping up to 1 million verifications per hour

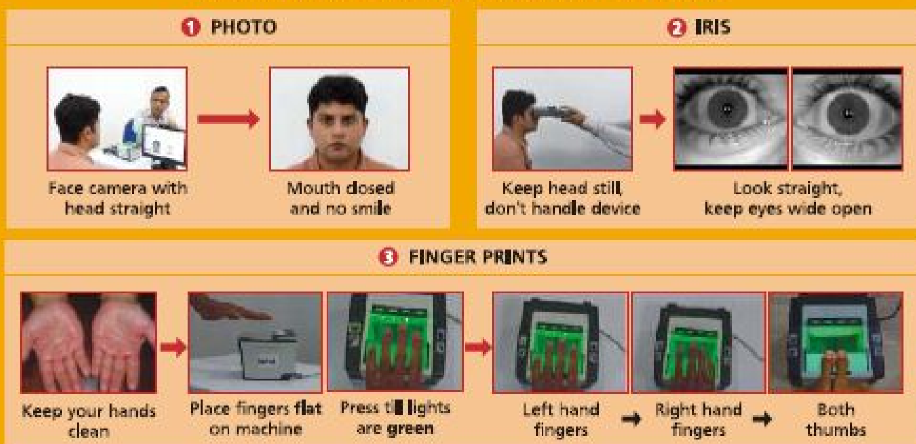
Verification must be accurate and quick

Unique Identification Authority of India
Planning Commission, Government of India

How to give biometrics for AADHAAR?

"We gave our biometric data to get our Aadhaar. Let us tell you how."

Biometric Process as part of Aadhaar Enrolment



Your facial photo, iris image of both the eyes and ten fingerprints will be scanned. Make sure you give the best quality biometric data possible, as it will be required to generate your Aadhaar number and to prove your identity through Aadhaar in the future. Enrolment can happen even with any permanent or temporary handicap of fingers or eyes.

You can Enrol Anywhere in India | Aadhaar Enrolment is Free of Cost | You need to Enrol Only Once

www.uidai.gov.in

Watch this space tomorrow to know "What happens after Aadhaar Enrolment"

AADHAAR
Aam Aadmi Ka Adhikaar

BEST FINGER DETECTION (BFD)

Different technologies and processes used at enrollment and at verification

During first verification, measure similarity score between new image and enrollment image for all 10 fingers

→ Determine “best” finger (BFD)

Use BFD information in all subsequent verifications

Using best finger performs much better than using right thumb

OUR STUDY

Goal: Derive verification policy that is accurate and does not impose too much delay

Approach:

- New parametric biometric model for 12 measurements

- Calibrate model using UIDAI experiments

- Introduce Best Iris Detection (BID) process (like BFD)

- Find near-optimal individualized policies

  Different subset of 12 biometrics used for each person

- Compare performance to policies considered by UIDAI

FINGERPRINT MODEL

X_i = true (vs. measured) genuine (vs. imposter) similarity score between enrollment image and verification image for finger $i = 1, \ldots, 10$

Each resident has overall image quality $\theta_i \sim \mathcal{N}(\mu, \tau^2)$
where μ is overall mean
τ^2 is interperson variance

Given realization of θ, $\ln X_t \sim \mathcal{N}(c_i\theta, \sigma^2)$
where c_i is finger-dependent correction ($\sum_i c_i = 10$)
σ^2 is intraperson interfinger variance

$\ln Y_i = \ln X_i + \delta_i$ = score during BFD process

$\ln Z_i = \ln X_i + \epsilon_i$ = score during subsequent verifications
where δ_i and ϵ_i are measurement noise

Iris similarity scores are bivariate lognormal plus measurement noise

PROPOSED SINGLE-STAGE POLICY

Observe BFD and BID scores ($Y_1, \ldots, Y_{12}$)

Choose subset of 12 biometrics (10 fingers + 2 irises)

Choose threshold

PROPOSED SINGLE-STAGE POLICY

Observe BFD and BID scores ($Y_1, \ldots, Y_{12}$)

Choose subset of 12 biometrics (10 fingers + 2 irises)

Choose threshold

Observe new similarity scores for chosen subset (Z)

Compute likelihood ratio L (Prabhakar + Jain 2002)

$$L = \frac{P(\text{observe Z scores}|\text{genuine})}{P(\text{observe Z scores}|\text{imposter})}$$

Accept if L > threshold
Reject if L < threshold

APPROACH

Challenge: solve problem accurately and quickly

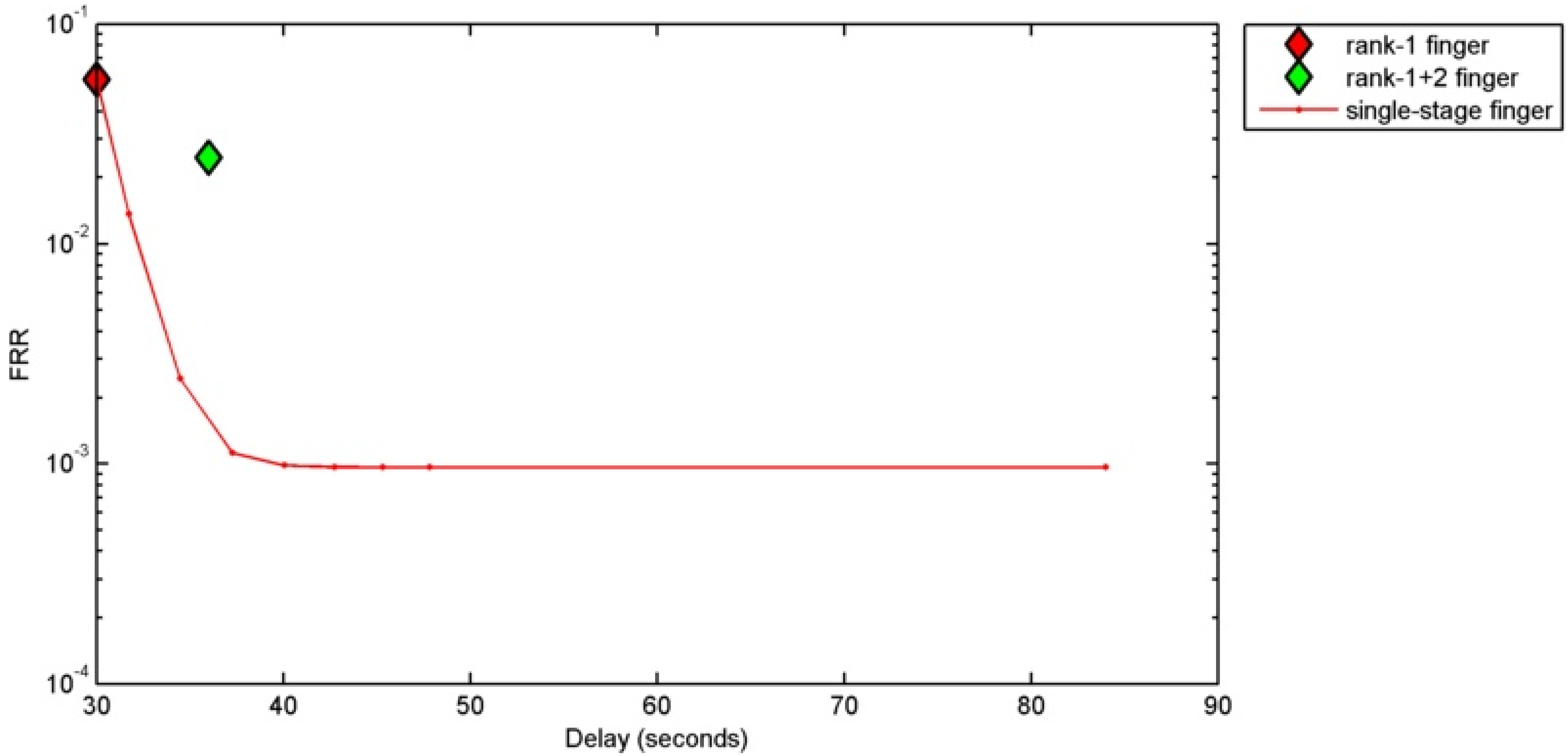
Approach: parametric model and analytical approximations

Fingers ranked by $\mu_i = c_i + 3.1 \ln Y_i$

$$c_i \approx 1\,, \;\; \ln Y_i \approx 4.1$$

Tools:
Clark's algorithm for maximum of normals
Conditional expectations of normals
Girsanov's change-of-measure theorem
Cubature (evaluating multi-dimensional surface integrals)
Gauss-Hermite quadrature for integrating normal PDFs
Gauss-Laguerre quadrature for integrating normal PDFs
Definite quadratic forms of normals = combo of non-central chi-squares
Approximation for non-central chi-square (Provost+Rudiuk 1996)

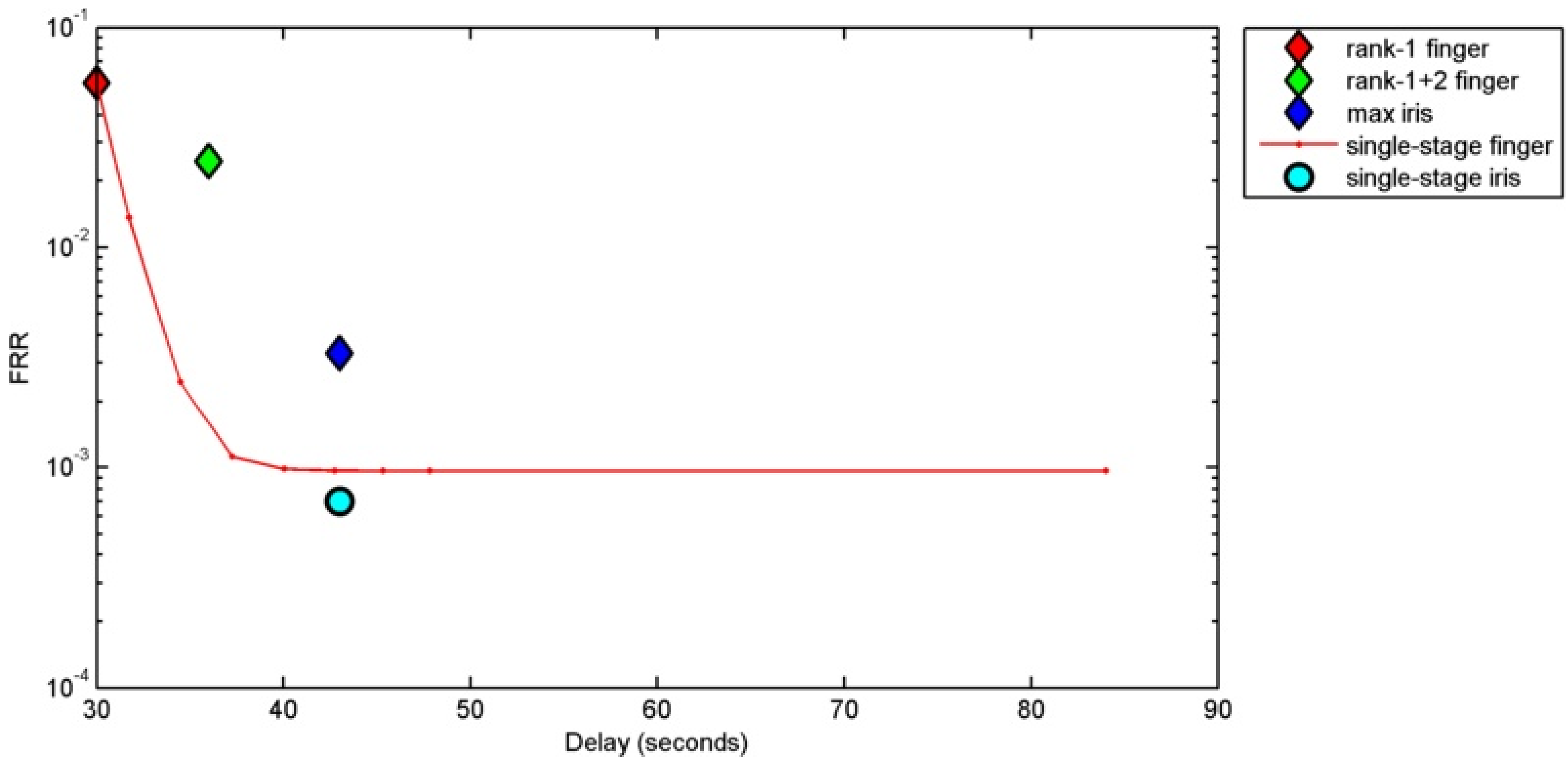
FRR vs. Delay
FAR=10^{-4}



Single-stage finger policy's FRR falls by 1.7 logs when delay increases from 30 sec (1 finger) to 40 sec (2.7 fingers)

Single-stage finger policy reduces FRR by 1.2 logs relative to sum of rank-1 and rank-2 fingers

FRR vs. Delay
FAR=10^{-4}



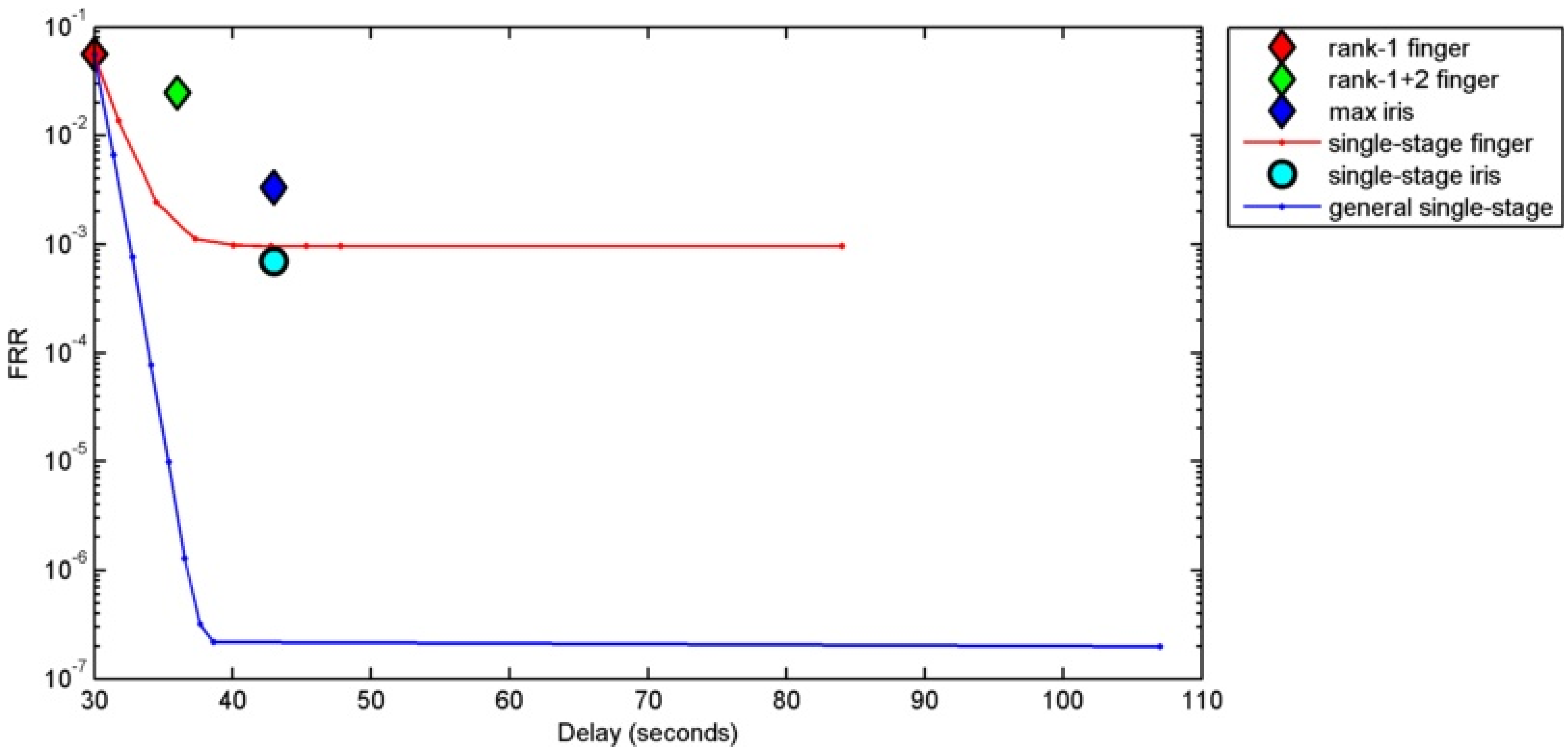
Single-stage iris policy reduces FRR by 0.7 logs relative to max iris policy

- due to use of likelihood ratio and both irises

Single-stage iris policy ≈ single-stage finger policy

FRR vs. Delay

FAR=10^{-4}



General single-stage policy:

- Gives 3.7 logs (5000-fold) reduction in FRR relative to single-stage finger policy for delay $\geq$ 38 sec.
- Acquires: irises from 37% of residents
  1.3 fingers on average
  both fingerprints and irises from 1-2% of residents

CONCLUSION

Optimal single-stage policy (delay = 38 sec) achieves same FRR as if all 12 biometrics were acquired from every resident (delay = 107 sec)

Optimal single-stage policy acquires:

- Iris scans from 32-41% of residents
- An average of 1.3 fingerprints per resident
- Both irises and fingerprints from only 1-2% of residents

Dramatic FRR reduction:

- 100,000-fold reduction vs. UIDAI's rank-1 + rank-2 finger policy
- 20,000-fold reduction vs. UIDAI's max iris policy (FAR $\geq 10^{-4}$)
- 5,000-fold reduction vs. UIDAI's max iris policy (FAR $\leq 10^{-5}$)

Two-stage policies only help when marginal delay cost between 30 and 37 seconds is very high

AFTERMATH

Cabinet level minister (Nandan Nilekani, Infosys) was briefed on our paper

Met with UIDAI Chief of Biometrics

Wrote briefing for India Supreme Court

They switched to two-stage finger policy