Biostatistics in Humanitarian Settings

Professor Miguel Angel Luque Fernandez, MA, MSc, GStat, PhD

Department of Statistics and Operations Research, University of Granada, Spain

March 1, 2024



UNIVERSIDAD DE GRANADA

Introduction

Biometrical work in resource-constraint settings comes with several challenges:

- Limited resources and time
- Idenfiying the optimal statistical method to answer a complex research question
- Evaluate the feasibility of interventions
- Evaluate the effectiveness of interventions
- Engaging with local partners: communication, empowerment and resources.

Overall objective

I would like to share three empirical examples where advanced biostatistical methods were essential to answering complex population-based research questions in resource constrain settings and humanitarian settings.

Then, I would like to reflect and briefly discuss the implications of the presented work.

Specific objective 1

I will share the **"club of patients"** research project where we used inverse probability weighting of marginal structural models to evaluate the effectiveness of a non-randomized intervention. [Luque-Fernandez et al., 2013]

M. A. Luque-Fernandez, G. Van Cutsem, E. Goemaere, K. Hilderbrand, M. Schomaker, N. Mantangana, S. Mathee, V. Dubula, N. Ford, M. A. Hernan, and A. Boulle. Effectiveness of patient adherence groups as a model of care for stable patients on antiretroviral therapy in khayelitsha, cape town, south africa. PLoS One, 8(2):e56088, 2013. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0056088.

https://www.ncbi.nlm.nih.gov/pubmed/23418518

Specific objective 2

I will show how I used classical biostatistical methods to improve the diagnostic accuracy of the mid-upper-circumference-arm (MUAC), a device tool, used to identify severe malnutrition in humanitarian crises. [Fernandez et al., 2010]

M. A. Fernandez, P. Delchevalerie, and M. Van Herp. Accuracy of muac in the detection of severe wasting with the new who growth standards. Pediatrics, 126(1):e195–201, 2010. ISSN 1098-4275 (Electronic) 0031-4005 (Linking). doi: 10.1542/peds.2009-2175.

https://www.ncbi.nlm.nih.gov/pubmed/20587675

https://scholar.harvard.edu/files/malf/files/pediatrics-2010-fernandez-e195-201.pdf

Specific objective 3

I will describe the reemergence of **cholera epidemics** linked to climatic change, evaluated by using auto-regressive Poisson time series analysis.[Luque Fernandez et al., 2009, 2012, 2011]

M. A. Luque Fernandez, A. Bauernfeind, J. D. Jimenez, C. L. Gil, N. El Omeiri, and D. H. Guibert. Influence of temperature and rainfall on the evolution of cholera epidemics in lusaka, zambia, 2003-2006: analysis of a time series. Trans R Soc Trop Med Hyg, 103(2):137–43, 2009. ISSN 1878-3503

https://doi.org/10.1016/j.trstmh.2008.07.017

Example 1: HIV ART Adherence Club of Patients

Club of patients as Model of Care for HIV patients



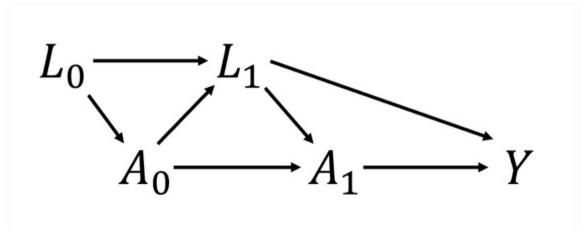
HIV: Club of Patients

- Piloted by Médecins Sans Frontières/Doctors Without Borders (MSF) in Khayelitsha, the award-winning ART Adherence Club model focuses on patient participation and peer support, for improved treatment adherence.
- This simple model allows patient groups to collect pre-packed, two-month supplies of treatment from lay health workers either at the clinic or outside of the clinic - whether at a local library or at a fellow patient's home.
- Today, the Cape Town Metro is home to more than 400 of these clubs as part of a
 partnership between MSF, the Western Cape Department of Health, Cape Town's City
 Health and the Institute for Healthcare Improvement.
- $\bullet \ https://www.msf.org.za/sites/default/files/art_adherence-club_report_toolkit.pdf \\$

HIV: Club of Patients, the challenge

- Originally a non-randomized intervention shifting tasks from doctors to lay workers to improve the adherence to the HIV treatment
- How to evaluate the effectiveness of non-randomized interventions?
- How to control for time-dependent confounders in longitudinal time-to-event settings?

DAG: Effect of club participation on LTF or DEATH



A=intervention (club participation), L = CD4, Y = Loss to follow-up and death.

HIV: Club of Patients

- We developed a retrospective observational evaluation of adherence clubs. Accordingly, we built a Marginal Structural Model (MSM) using Inverse Probability of Treatment Weighting (IPTW) estimator to estimate the intention-to-treat effect of adherence club participation.
- In the weighted analysis, club participation at any time following the start of the study
 was rendered independent of measured potential confounders by multiplying the
 propensity scores for starting or not starting club participation over time

Methods: IPTW for longitudinal data

Standardized weights in a longitudinal setting are estimated as

$$sw_{ij} = \prod_{k=0}^{j} \frac{P(A_{ik} = a_{ik} | \bar{A}_{ik-1} = \bar{a}_{ik-1}, V_i = v_i)}{P(A_{ik} = a_{ik} | \bar{A}_{ik-1} = \bar{a}_{ik-1}, \overline{C}_{ik} = \bar{c}_{ik}, V_i = v_i)}$$

where A is the exposure for subject i at time t_-ij (time points range starting at k=0 to k=j). The numerator contains the probability of the observed exposure at each time point (a_-ik) conditioned on the observed exposure history of the previous time point (a_-ik-1) and the observed non-time varying covariates (v_-i) . The denominator contains the probability of the observed exposure at each time point conditioned on the observed exposure history of the previous time point (a_-ik-1) , the observed time-varying covariates history at the current time point (c_-ik) , and the non-time varying covariates (v_-i) .



HIV: Club of Patients, the results

LTF or death

Unweighted model, no covariates HR= 0.23 (0.14-0.37)

Unweighted model, baseline covariates HR= 0.46 (0.26-0.82)

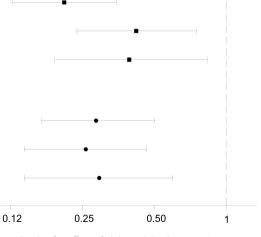
Weighted model, baseline covariates HR= 0.43 (0.21-0.91)

Virologic rebound

Unweighted model, no covariates HR= 0.32 (0.19-0.56)

Unweighted model, baseline covariates HR= 0.28 (0.16-0.52)

Weighted model, baseline covariates HR= 0.33 (0.16-0.67)

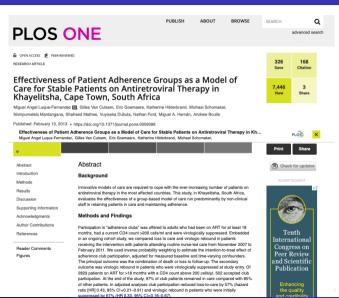


Hazard ratios for effect of club participation on outcomes

HIV: Club of Patients

- Patient adherence groups were found to be an **effective model** for improving retention and documented virologic suppression for stable patients in long term ART care.
- Out-of-clinic group-based models facilitated by non-clinical staff are a promising approach to assist in the long-term management of people on ART in high burden low or middle-income settings.

HIV: Club of Patients, the evaluation 1



HIV: Club of Patients, conclusion

- We need robust statistical methods to prove evidence and change policy
- We need to disseminate the results and findings outside academia
- Partnerships between academia and other agencies (i.e., non-governmental or international) is key
- Dissemination and democratization of advanced statistical methods is paramount (i.e., open source available tutorials)



Introduction to computational causal inference using reproducible Stata, R, and Python code: A tutorial

Abstract

The main purpose of many medical studies is to estimate the effects of a treatment or exposure on an outcome. However, it is not always possible to randomize the study participants to a particular treatment, therefore observational study designs may be used. There are major challenges with observational studies; one of which is

Targeted Maximum Likelihood Estimation for a Binary Treatment: A tutorial. Statistics in Medicine, 2018

This repository made available to the scientific community the data and the code presented in the Statistics and Medicine (SIM) manuscript:

Luque-Fernandez, MA, Schomaker, M, Rachet, B, Schnitzer, ME. Targeted maximum likelihood estimation for a binary treatment: A tutorial. Statistics in Medicine. 2018; 37: 2530–2546. https://doi.org/10.1002/sim.7628

Link to the Statistics in Medicine Open Access Article



Miguel Angel Luque Fernandez, Michael Schomaker, Bernard Rachet, Mireille Schnitzer

ABSTRACT

When estimating the average effect of a binary treatment (or exposure), methods that incorporate propensity scores, the G-formula, or targeted maximum likelifond estimation (TMLE) are preferred over naive regression approaches which are biased under misspecification of a parametric outcome model. Contrastingly, propensity score methods require the correct specification of an exposure model. Double-robust methods only require correct specification of one of these models. TMLE is a semi-parametric double-robust method that improves the chances of correct model specification by allowing for flexible estimation using non-parametric machine-learning methods. It therefore requires weaker assumptions than its competitions. We provide a step-by-step guided implementation of TMLE and illustrate it in a realistic scenario based on cancer epidemiology where assumptions about correct model specification and positivity (i.e., when a study participant had zero probability of receiving the treatment) are nearly violated. This article provides a concise and reproducible educational introduction to TMLE for a binary outcome and exposure. The reader should gain sufficient understanding of TMLE from this introductory turinal to be able to apply the method in practice. Extensive R-code is provided in easy-to-read boxes throughout the article for replicability. Statu suers will find a testing implementation of TMLE and additional partical in the amongram status of the following Gift-the prositions that incomplementation of TMLE and additional field of the following Gift-the prositions that the provides in asset in the following Gift-the prositions that the provides in asset in the following Gift-the prositions that the provides in asset in the following Gift-the prositions that the provides in asset in the following Gift-the provides in asset the contract of the following Gift-the provides in asset in the particle provides in asset in the following Gift-the provides in asset in the following Gift-the provides in asset in the follow

International Journal of **Epidemiology**



Issues

More Content •

Submit *

Purchase Alerts

About •

International Journal of Er ▼ Search





ent



Volume 48, Issue 2 April 2019

Article Contents

Abstract

JOURNAL ARTICLE

Educational Note: Paradoxical collider effect in the analysis of non-communicable disease epidemiological data: a reproducible illustration and web application

Miguel Angel Luque-Fernandez 록. Michael Schomaker, Daniel Redondo-Sanchez. Maria Jose Sanchez Perez, Anand Vaidva, Mireille E Schnitzer

International Journal of Epidemiology, Volume 48, Issue 2, April 2019, Pages 640-653. https://doi.org/10.1093/iie/dvv275

Published: 14 December 2018 Article history ▼

A correction has been published: International Journal of Epidemiology, Volume 49. Issue 1, February 2020, Page 356, https://doi.org/10.1093/ije/dyz247

A correction has been published: International Journal of Epidemiology, Volume 50. Issue 2, April 2021, Page 699, https://doi.org/10.1093/ije/dyab006











ALTMETRIC





the Simons Foundation, Colum

Search...

ITX1V > stat > arXiv:2206.15310

Statistics > Methodology

[Submitted on 30 Jun 2022]

The Delta-Method and Influence Function in Medical Statistics: a Reproducible Tutorial

Rodrigo Zepeda-Tello, Michael Schomaker, Camille Maringe, Matthew J. Smith, Aurelien Belot, Bernard Rachet, Mireille E. Schnitzer, Miguel Angel Luque-Fernandez

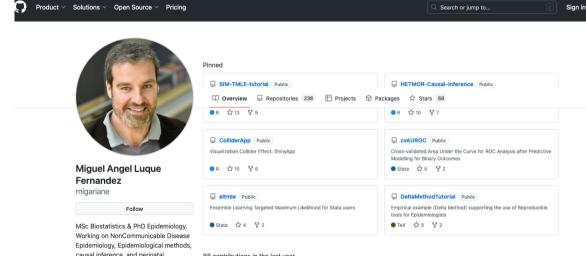
Approximate statistical inference via determination of the asymptotic distribution of a statistic is routinely used for inference in applied medical statistics (e.g. to estimate the standard error of the marginal or conditional risk ratio). One method for variance estimation is the classical Delta-method but there is a knowledge gap as this method is not routinely included in training for applied medical statistics and its uses are not widely understood. Given that a smooth furnition of an asymptotically normall estimator is also asymptotically normally distributed, the Delta-method allows approximating the large-sample variance of a function of an estimator with known large-sample properties. In a more general setting, it is a technique for approximating the variance of a functional (h.e., an estimand) that takes a function as an input and applies another function to it (e.g. the expectation function). Specifically, we may approximate the variance of the function using the functional Delta-method based on the influence function (IP). The IF explores how a functional $\phi(\theta)$ changes in response to small perturbations in the sample distribution of the estimator and allows computing the empirical standard error of the distribution of the functional. The ongoing development of new methods and techniques may pose a challenge for applied statisticians who are interested in mastering the application of these methods. In this tutorial, we review the use of the classical and functional Delta-method and their links to the IF from a practical perspective. We illustrate the methods using a cancer epidemiology example and we provide reproducible and commented code in R and Python using symbolic programming. The code can be accessed at this https URL

Subjects: Methodology (stat.ME)

Cite as: arXiv:2206.15310 [stat.ME]

(or arXiv:2206.15310v1 [stat.ME] for this version) https://doi.org/10.48550/arXiv.2206.15310

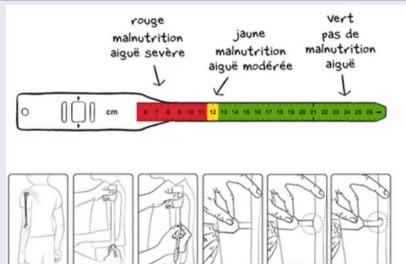




https://github.com/migariane

Example 2: MUAC

Mid-upper-arm-circumference (MUAC) measuring tape



MUAC: The device

Mid-upper-arm-circumference arm is a deviced used in humanitarian settings to detect malnutrition in children aged between 6 months and 5 years. It is a simply bracelet that is wrapped around the arm that uses a colour code to determine the degree of malnutrition:

- red for severe acute malnutrition,
- yellow for moderate acute malnutrition, and
- green for no acute malnutrition.

MUAC: WHAT'S KNOWN ON THIS SUBJECT

- MUAC measurements are used to screen rapidly for malnutrition among children 6 to 59 months of age.
- With the introduction of a new growth curve for children by the WHO in 2006, an evaluation of MUAC diagnostic accuracy is needed.

MUAC: OBJECTIVE

To estimate the **accuracy** of using mid-upper-arm circumference (MUAC) measurements to diagnose severe wasting by comparing the new standards from the World Health Organization (WHO) with those from the US National Center for Health Statistics (NCHS).

MUAC: WHAT THIS STUDY ADDS

- This study confirms the need to change the MUAC cutoff value from 110 mm to 115 mm.
- This change is needed to maintain the same diagnostic accuracy and to identify children at greatest risk of death resulting from severe wasting.

MUAC: Classical Biostatistics Methods

- Statistical analyses: we calculated the nutritional indicators of severe wasting, weight-for-height z scores for 34,937 children between the ages of 6 and 59 months, from 39 nutritional surveys conducted by Doctors Without Borders across the world, according to the NCHS and WHO curves.
- We used 2x2 tables to determine the sensitivity, specificity, positive predictive value, and Youden index of various MUAC cutoff points (110, 115, 125, 135, 140, and 145 mm).
- We used ROC curves to estimate the AUC for different MUAC cutoff values, comparing the discriminatory capacity of the WHO and NCHS curves for severe wasting.

MUAC: RESULTS

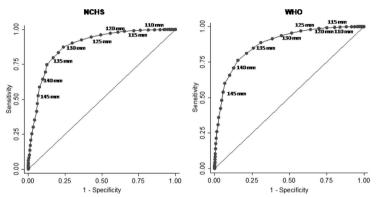


FIGURE 1

Receiver operating characteristic curves for severe wasting, defined as weight-for-height z scores below —3, with NCHS (area under the curve: 0.82 [95% CE: 0.79—0.821) and WHO (area under the curve: 0.80 [95% CE: 0.79—0.821) standards and different MUAC cutoff values.

Figure 2: Source: https://www.ncbi.nlm.nih.gov/pubmed/20587675



Technical Bulletin No.13 revision 2

MID-UPPER ARM CIRCUMFERENCE (MUAC) MEASURING TAPES

Background:

A range of Mid-Upper Arm Circumference (MUAC) Measuring Tapes are available through UNICEF Supply Division. MUAC tapes are predominately used to measure the upper arm circumference of children but also that of pregnant women, helping identify mainutrition.

There are different types of MUAC tape available. All are graduated in millimetres and some are colour coded (red, yellow and green) to indicate the nutritional status of a child or adult. The colour codes and gradations vary depending on the tape type.

In May 2009, the Word Health Organization (WHO) and UNICEF issued a joint statement on WHO child growth standards and the identification of severe acute malnutrition in infants and children. To reflect this, a new standard MUAC tape (50145620 MUAC, Child 11.5 Red/PAC-50) was made available.

Before ordering, you should check if the government in your country has implemented the new WHO standards, in which case order **S0145620**. If your government has not implemented the new standards, order **S0145600**.

The item description text below has been changed to reflect the adjusted tape cut-off points. MUAC tapes for adults are also now available through Supply Division warehouse. We have also included the beneficiary type (child, adult) into the item descriptions.

MUAC tapes available through Supply Division

S0145620 MUAC, Child 11.5 Red/PAC-50

This is a new item. It was created in order to support implementation of the new standards (see above).

Cut-off points of S0145620:

Red:	0 – 11.5 cm				
Yellow:	11.5 cm - 12.5 cm				
Green:	from 12.5 cm				

MUAC: UNICEF CHANGE OF POLICY

 According to UNICEF guidelines, children with a MUAC of less than 11.5 cm, have a higher risk of death.

MUAC: DISCUSSION

- Change of international policy is challenging
- We need large and representative sample sizes
- Classical biostatistical methods are robust
- Public health evidence requires the appropriate use of robust and consolidated statistical methods.

Example 3: Cholera

Reemergence of Cholera Epidemics and Climatic Change

A failure to slow global warming is providing many deadly diseases with the opportunity to expand their reach, putting the health of millions of people at risk. Read on to understand how climate change and infectious disease are linked, and what we can do to limit the damage.



Figure 4: Credit: Alexis Huguet: All Rights Reserved

Cholera: Evidence in 2021





Search Here...

About the website

Health threats due to Climate ~ Change

Climatic trends and variability

Climate change impacts on disasters

Climatic factors on local scale

Climatic impacts on global and regional scale Responses to global warming and ~ climate change

Publications >

Cholera

Home > Health threats due to Climate Change > Water pollution and water borne diseases > Cholera

Climate Change And Cholera

Cholera is an acute diarrheal disease caused by the bacterium Vibrio cholerae (V. cholerae). Cholera incidence has been greatly reduced due to improved environmental conditions and the implementation of intervention measures [1, 2]. However, it remains a global threat to public health and has emerged in some areas [2]. It is estimated that cholera cases range from 1.3 million to 4.0 million each year worldwide, resulting in 21,000 to 143,000 deaths [3]. Evidence suggests that climate

Useful Links

IEDCR

icddr,b

Cholera: Evidence to support the link was needed







Influence of temperature and rainfall on the evolution of cholera epidemics in Lusaka, Zambia, 2003–2006: analysis of a time series

Miguel Ángel Luque Fernández^{a,*}, Ariane Bauernfeind^b, Julio Díaz Jiménez^c, Cristina Linares Gil^a, Nathalie El Omeiri^{a,d}, Dionisio Herrera Guibert^e



National Centre of Epidemiology (CNE), Programa de Epidemiologia Aplicada de Campo (PEAC), Instituto de Salud Carlos III (ISCIII), C/Sinesio Delgado 6, Pabellón 12, 28029 Madrid, Spain

^b Médecins Sans Frontières (MSF), Brussels Operational Centre, Rue Dupré 94, 1090 Bruxelles, Belgium

c Escuela Nacional de Sanidad (ENS), Instituto de Salud Carlos III (ISCIII), C/Sinesio Delgado 6, Pabellón 7, 28029 Madrid, Spain

^d European Programme for Intervention Epidemiology Training (EPIET), Stockholm, Sweden

[°] National Centre of Epidemiology (CNE), Instituto de Salud Carlos III (ISCIII), C/Sinesio Delgado 6, Pabellón 12, 28029 Madrid. Spain

Cholera: Methods

- Poisson autoregressive model controlling for seasonality.
- Standard errors of coefficients were scaled to control for oversdispersion.

Cholera: Methods

- Sine and cosine functions were used in the model for building the independent variables that explain the seasonal component of the series.
- Based on our review of the literature, an autoregressive term at order 1 was incorporated into the model to control for the autocorrelation, and lags of up to 8 weeks for temperature and rainfall.

Cholera: Results

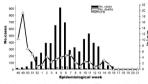


Figure 1 Distribution of cholera cases, deaths and case fatality rate (CFR) per epidemiological week (n=6471 cases and 205 deaths) in the 2003–2004 outbreak in Lusaka, Zambia (Médecins Sans Frontières, unpublished data).

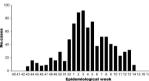
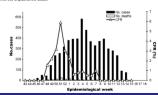


Figure 2 Distribution of cholera cases per epidemiological week (n=888 cases) in the 2004—2005 outbreak in Lusaka, Zambia (Médecins Sans Frontières, unpublished data).



Analysis of the evolution of cholera epidemics in Zambia

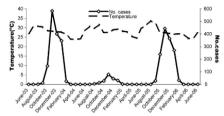
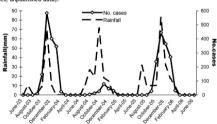


Figure 4 Time plots of number of cholera cases per month and monthly mean temperature (°C) in Lusaka, Zambia, 2003–2006 (Médecins Sans Frontières, unpublished data).



38 / 43

Cholera: Results

Table 2 Association between the number of cholera cases and climate variables: final autoregressive Poisson model including lags of weekly mean temperature and rainfall (Médecins Sans Frontières, unpublished data)^a

	Coefficient (SE) ^b	RR (95% CI)	% change ^c	AR (%)	<i>P</i> -value
Temperature (6 weeks earlier)	0.05 (0.006)	1.05 (1.04-1.06)	5.2	4.7	<0.001
Rainfall (3 weeks earlier)	0.02 (0.01)	1.02 (1.01-1.04)	2.5	1.9	0.011

RR: relative risk; AR: attributable risk.

Please cite this article in press as: Luque Fernández MÁ, et al. Influence of temperature and rainfall on the evolution of cholera epidemics in Lusaka, Zambia, 2003–2006: analysis of a time series. *Trans R Soc Trop Med Hyg* (2008), doi:10.1016/j.trstmh.2008.07.017

Figure 9: MA Luque-Fernandez

Adjusted for seasonality.

^b Standard errors (SE) scaled using square root of Pearson χ^2 based dispersion.

^c Percent change in expected count for 1 °C increase in temperature and 50 mm in rainfall.

Cholera: Conclusion

I found an association between an increase in the number of cholera cases and climate variables:

6 weeks prior to the beginning of the rainy season an increase in temperature is observed followed by an increase in rainfall 3 weeks later, both exceeding expected levels, we may be confronted with an increase in the number of cases of cholera within the following 3 weeks.

Cholera: Discussion

- Applied mathematical statistics is key to answer complex research questions
- Biostatisticians and Statisticians in limited resource settings need a wide range of expertise and practice in statistical methods and epidemiology.
- Published evidence is key to change policy

Overall Discussion

- More Biostatistics/Statistics and epidemiology in Humanitarian settings is needed
- A strong methodological background is key
- The variation and complexity of research questions requires of expertise, motivation, ambition and compromise
- Collaborative work and communication is paramount
- A decolonizing perspective is much needed and strong partnerships between academic and non academic institutions across the world is needed
- The bidirectional transfer of knowledge and capabilities is rewarding
- Biostatistical methods are fun

Thank you



