

# Death of children with SAM diagnosed by WHZ or MUAC: Who are we missing?

## Summary of presentation<sup>1</sup>

By Michael H. Golden and Emmanuel Grellety



Michael Golden is a retired professor of medicine with 45 years' experience of studying all aspects of malnutrition.



Emmanuel Grellety has spent his whole working life with humanitarian organisations in many roles. He is now an epidemiologist working with Epicentre and is completing his PhD.

We would like to thank Action Against Hunger for the opportunity to present this extended abstract to the R4NUT conference, to the editors of ENN and the reviewers of our submitted papers for very helpful comments on initial submissions, and to the agencies that provided patient data for our empirical study.

### Location: *Global*

**What we know:** Both weight-for-height z-score (WHZ) and mid-upper arm circumference (MUAC) are recommended to identify severely malnourished children for treatment. MUAC has distinct advantages for community-level screening; however several countries have gone further to instigate MUAC-only admissions for treatment.

**What this article adds:** A recent review examined the consequences of excluding children with severe acute malnutrition (SAM) identified using WHZ from admission to treatment programmes. Analysis of individual data from 14,935 children admitted to a range of treatment programmes over 22 years and a literature review examined case fatality rates (CFR) with different indicators and caseload. Simpson's paradox (mathematical coupling) results in reversal of significance that affects interpretation of the relative mortality rates of WHZ and MUAC. The analysis suggests that children with SAM identified by WHZ <math><-3Z</math> and admitted for treatment have as high a risk of death as children in treatment with MUAC <math><115\text{mm}</math>. Review of 21 datasets that compared WHZ and MUAC mortality rates show problems with interpretation of the reported CFRs; inconsistencies greatly limit analysis, comparability and interpretation. Caseload is a more important determinant of the number of SAM-related child deaths than the relative CFR to give SAM-attributable deaths. Where most of the children are identified as SAM using WHZ rather than MUAC, it is estimated that fewer than half of all SAM-related deaths will be identified using a MUAC-only programme. Strong advocacy for the use of MUAC to maximise coverage of treatment programmes has developed into MUAC-only programmes that are inadequately evidenced on the consequences of excluding WHZ cases. Urgent research is needed to develop simple methods to identify children with low WHZ at community level.

### Introduction

About 19 million children are estimated to have severe wasting, of whom about half to one million die each year (Black et al, 2013). These estimates were made using only weight-for-height z-score (WHZ) as the diagnostic criterion. As deaths related to a low mid-upper arm circumference (MUAC) were not taken into consideration, the actual number of deaths is much higher. This is because, of all the children with severe acute malnutrition (SAM), only about 16.5 per cent have both a WHZ (<math><-3Z</math>) and a MUAC (<math><115\text{mm}</math>) below the World Health Organization (WHO) defined criteria for SAM; the remainder have SAM by either WHZ (45 per cent) or MUAC (39 per cent), but not both criteria (Grellety and Golden, 2016); the degree of overlap varies greatly by context. Based on these figures, a MUAC-only programme would identify 55 per cent of all SAM children and a WHZ-only programme 61 per cent. Although the risks of death may differ from place to place and time to time, the actual number of SAM related deaths depends on the relative number of children fulfilling each criterion in the community, as well as the case fatality rates (CFR); that is, both the relative caseloads and mortality risks combine to give the total number of deaths occurring due to SAM.

Because of its simplicity, ease of use and cheapness, absolute MUAC has been readily taken up to screen for children with SAM in the community. WHO guidelines recommend the use of MUAC (and examination for bilateral pitting oedema) in children 6-59 months of age at community level for early identification and referral of children with

SAM for full assessment at a treatment centre (admission is then by MUAC or WHZ). However, many agencies and several national governments (e.g. Nigeria, South Sudan, Bangladesh) have gone further and ceased attempting to identify and treat any children with SAM diagnosed by WHZ. They have based this upon repeated advocacy for MUAC-only programmes, justified by its simplicity, and reports purporting to show a universally higher mortality risk for SAM identified by MUAC (SAM-muac) than SAM identified by WHZ (SAM-whz). The latter is based largely upon statistical comparison of ROC curves and the conclusion that children with SAM diagnosed by WHZ, but not by MUAC, are at lower risk of death (en-net, 2015a; en-net, 2015b; Briend et al, 2016). Our serious concerns regarding the consequences of excluding SAM-whz children from admission to treatment programmes prompted this review, where we examine the relative mortality rates from a large number of SAM children; appraise the literature with consideration of the statistics and methods used; and analyse the numbers of deaths likely to be missed if a MUAC-only policy were to be universally adopted.

### Methods

We obtained individual data from 14,935 children treated in inpatient facilities (IPF), 45,364 treated in outpatient treatment programmes (OTP) and 16,588 patients initially admitted to supplementary feeding programmes (SFPs) as moderately malnourished but who, with the change

<sup>1</sup> Presentation at the ACF Research for Nutrition Conference, Pavillon de L'Eau, 13th November, 2017. A video of the presentation can be found here: [https://youtu.be/yIWjGG\\_S5YU](https://youtu.be/yIWjGG_S5YU)

in diagnostic cut-off points and standards, would now be reclassified as SAM. WHO 2006 criteria and the presence or absence of oedema were used to divide the children into seven groups, depending upon the various combinations of diagnostic criteria (see footnote to Table 1 for explanation of six groups, plus kwashiorkor not shown). We conducted an exhaustive search of the literature to identify reports of children diagnosed by WHZ or MUAC with the respective mortality rates. The papers were reviewed. We analysed the effect of caseload using published prevalence data and CFRs derived from our empirical data, the literature data and theoretical simulations.

## Results

### Empirical data

Table 1 shows the CFR of all the patients with SAM by diagnostic category. The CFR was higher for those with marasmic SAM admitted with only a low WHZ than for those with only a low MUAC. The children who had both diagnostic criteria had a significantly higher CFR. When the children fulfilling both criteria are included in each of the diagnostic groups, the relative CFR of children admitted by WHZ vs MUAC is reversed, so that it now appears that MUAC-associated mortality is higher than with WHZ. This is an example of Simpson's paradox (see illustration in Table 2), caused in this case by mathematical coupling (Tu et al, Archie Jr, 1981). Oedematous children

who had a low WHZ had a much higher CFR than those with a low MUAC; all CFRs were higher for SAM children with oedema than without oedema. Although the relative mortality was not quite reversed when the children with both anthropometric deficits were considered, the difference was considerably ameliorated.

When all the children's data are considered together, WHZ-related death rate was higher than MUAC-related deaths; those who had both deficits had about twice the CFR as those with a single anthropometric deficit. When those with both a low WHZ and a low MUAC were included in the WHZ and MUAC data, the relative CFR is reversed. This demonstrates that inclusion of children with both deficits into both the WHZ and the MUAC group results in erroneous interpretation of the relative mortality rates.

### Literature review

We retrieved 21 datasets that compared WHZ and MUAC mortality rates. Table 3 shows the problems with the interpretation of the reported CFRs. Statistically, to get reliable results, the expected deaths in each group should be at least five. The reports marked in *brown* had insufficient deaths to make analyses of individual studies reliable. Most of the "brown" studies had many more children fulfilling the "both" category; as shown with the empirical data this also makes

the analyses reported by the authors subject to mathematical coupling and are thus unreliable. The reports marked in *pink* each suffer from the same criticism by including individual children with both SAM by WHZ and MUAC criteria into the MUAC and WHZ groups; the analyses are therefore flawed. Children fulfilling both criteria would be identified with all screening strategies and not excluded from treatment. The papers marked in *blue* included oedematous children in the analysis; as oedema greatly increases the mortality risk, this confounding further increases the unreliability of the reports. The *purple* columns indicate the papers which used obsolete standards. This affects the CFR because more stringent criteria include more severely affected children in the cohort resulting in a higher CFR; less stringent criteria have the opposite effect, so that a lower CFR is expected when less severely affected children are considered to have SAM.

Figure 1 shows how the criteria used for WHZ have changed. For example, where Centers for Disease Control and Prevention (CDC) 2000 criteria<sup>2</sup> are used, one expects a lower CFR than if WHO criteria are used. Similarly, where a MUAC cut-off of <110mm is used instead of the WHO recommended criterion of <115mm, a higher CFR is expected. There are three reports that use the same data.

<sup>2</sup> [https://www.cdc.gov/growthcharts/cdc\\_charts.htm](https://www.cdc.gov/growthcharts/cdc_charts.htm)

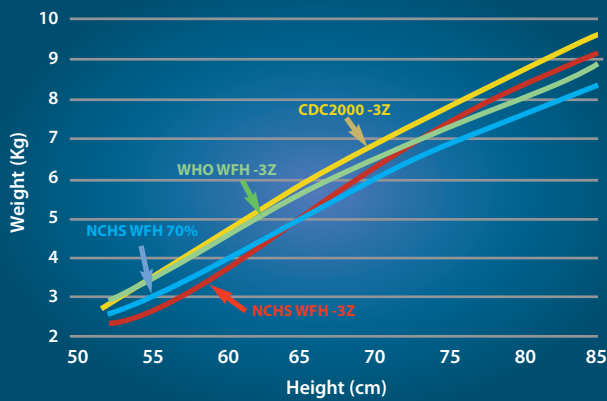
Sub-groups	Total #	Dead #	CFR %	Comparisons	p
<b>MUAC v WHZ without oedema</b>					
M-muac	7,191	56	0.78	muac v whz	0.000
M-whz	16,530	332	2.01	muac v both	0.000
M-both	40,307	2,000	4.96	whz v both	0.000
M-All-muac	47,498	2,056	4.33	$\chi^2=3.01$	0.083
M-All-whz	56,837	2,332	4.1		
<b>MUAC v WHZ with oedema</b>					
K-muac	1,669	118	7.07	muac v whz	0.000
K-whz	1,088	169	15.53	muac v both	0.000
K-both	4,217	576	13.66	whz v both	0.411
K-All-muac	5,886	694	11.79	$\chi^2=9.75$	0.002
K-All-whz	5,305	745	14.04		
<b>MUAC v WHZ with and without oedema</b>					
M+K-muac	8,860	174	1.96	muac v whz	0.000
M+K-whz	17,618	501	2.84	muac v both	0.000
M+K-both	44,524	2,576	5.79	whz v both	00.000
M+K - All-muac	53,384	2,750	5.15	$\chi^2=2.16$	0.141
M+K - All-whz	62,142	3,077	4.95		

WHZ refers to children with WHZ <-3Z and MUAC >115mm, MUAC refers to children with WHZ >-3Z and MUAC <115mm. "Both" refers to children with both a MUAC <115mm and WHZ <-3Z. The prefix "M" denotes marasmic children; prefix "K" denotes oedematous malnutrition (marasmic kwashiorkor). "M+K" indicates children with either marasmus or oedema. The prefix "All" indicates all the children who would be diagnosed with a MUAC-only programme; thus, "All-muac" is defined as M-muac + M-both (or K-muac + K-both) and "All-whz" is defined as M-whz + M-both (or K-whz + K-both).

	Total #	dead #	CFR %	CFR X/Y	X deaths %	Y deaths %
<b>Scenario A</b>						
X	500	0	0.0			
Y	1500	30	0.2	0.0		
Both X & Y	500	30	0.6			
Total	2,500	60	2.4			
All-X	1,000	30	0.3		50	
All-Y	2,000	60	0.3	1.0		100
<b>Scenario B</b>						
X	500	6	1.2			
Y	1500	30	0.2	0.6		
Both X & Y	500	30	0.6			
Total	2,500	66	2.6			
All-X	1,000	36	3.6		55	
All-Y	2,000	60	0.3	1.2		91
<b>Scenario C</b>						
X	500	10	0.2			
Y	1500	30	0.2	1.0		
Both X & Y	500	30	0.6			
Total	2,500	70	2.8			
All-X	1,000	40	0.4		57	
All-Y	2,000	60	0.3	1.3		86

Two variables, X and Y, are shown with various CFRs and caseloads. When subjects with both deficits are added to X and Y, the apparent CFR of X is ameliorated (A), reversed (B), or inappropriately shown to be superior (C); despite this, the total numbers of deaths that will be identified if X is used as the diagnostic parameter is always much less than with Y. Mathematical coupling always occurs where "one variable directly or indirectly contains the whole or part of another, and the two variables are analysed using standard statistical techniques." (Archie Jr, 1981)

**Figure 1** The cut-off points for weight-for-height using different reference criteria



A child's MUAC is measured in a nutritional centre in Kalonda, Kasai Region, DRC, 2017

Report 16 has data for deaths with MUAC <115mm and National Center for Health Statistics (NCHS) criteria which shows a much higher CFR with WHZ; in report 17 the MUAC criterion has been reduced to the more stringent MUAC <110mm and the less stringent CDC 2000 criteria for WHZ, resulting in a reversal of the interpretation of the data so that MUAC now appears to have a much higher mortality rate. Report nine used WHO criteria and there is a non-significant higher mortality in the

WHZ group. Reports that do not use the current standards cannot be appropriately interpreted and can give a biased impression when applied for identification of children by WHO standards.

The green columns did not include children in the usual age range. Two of the studies had extremely short average observation periods (<4 days), which raises the question of verification bias and confounding by acute illness unrelated to malnutrition (e.g. convulsions).

The papers are sufficiently problematic that they cannot be used to guide policy decisions. Statistical analysis is limited as outlined and most have included oedematous children, have too few events, used obsolete standards or had a combination of these defects, which makes them individually inadequate evidence on which to promote MUAC-only programmes. The results are in broad agreement with the empirical data. It is concluded that children with SAM by MUAC-

**Table 3** An overview of the literature comparing WHZ and MUAC SAM deaths

No	Dataset		dead	dead	dead	both	kwash	diag	diag	Age	time	MUAC	WHZ	Both	MUAC:WHZ		
			MUAC	WHZ	Both	incl	incl	MUAC (mm)	WHZ			Range (m)	d	CFR	CFR	CFR	χ <sup>2</sup>
			#	#	#									%	%	%	p
1	Aguayo 2015	IPF	2	3	27							0.23	0.23	0.70	ns		
2	Grellety 2012	com	2	3	8					6-23		1.20	1.69	4.30	ns		
3	Grellety 2015	OTP	1	13	26							4.76	0.87	3.72	ns		
4	Isanaka 2015	OTP	4	3	6							0.43	0.57	0.64	ns		
5	Lowlaavar 2016	IPF	0	11	10						3 d	0.00	6.43	16.95	ns		
6	LaCourse 2014	IPF	4	2	1			<110	<70% NCHS		ng	5.88	6.90	6.25	ns		
7	Berkley 2005	IPF					incl		<-3 NCHS	12-59	ng	10.94	10.11	25.36	ns		
8	Chiabi 2017	IPF	3	1	18		incl				ng	14.29	4.35	29.03	ns		
9	Olofin 2016	com					ng				ng	13.25	15.50	37.50	ns		
10	Sachdeva 2016	IPF					incl				3.7 d	20.00	7.89	18.06	0.011		
11	Burza 2016	com				both	ng					5.54	8.81		0		
12	Morgeni 2011	IPF				both	incl				ng	16.83	8.90		0.024		
13	Sylla 2015	IPF				both	ng			0-60	ng	7.74	23.26		0.014		
14	Vella 1990	com				both	incl		<-3 NCHS	0-60	ng	18.18	8.82		ns		
15	Dramaix 1993	IPF+Com				both	incl		<70%NCHS	0-60		33.06	20.22		ns		
16	Garenne 1987	com				both	incl		<-3 NCHS	1-60		8.30	16.98		ns		
17	Garenne 2009	com				both	incl	<110	CDC2000	1-60		21.50	12.42		ns		
18	Garenne 2009	com				both	incl	<110	CDC2000	1-60		6.61	6.10		ns		
19	Van Den Broeck 1993	com	6	3		both	incl	<-4 SD	<-3 NCHS	0-60		1.00	2.00		ns		
20	Savadogo 2007	IPF				both		Tertile	<-3 NCHS	0->35		24.50	18.03		0.031		
21	Girum 2017	IPF				both	incl		<70% ng	ng	13 d	14.80	20.00		ns		

References are available on request from the authors. Brown = <5 expected events; pink = both MUAC and WHZ included; blue = oedematous malnutrition included; purple = obsolete standards; green = not 6-59 months age range; orange = low observation period; CFR = case fatality rate; ng = not given; IPF = inpatient facility; com = community; OTP = outpatient therapeutic programme



**Table 4** Proportion of total SAM deaths identified with a MUAC-only programme with various CFRs

	Data			Theoretical			Empirical	Literature reports	
	A	B	C	All	1-6	7-10			
	CFR S-muac %	5	10	10	0.78	0.63	12.88		
	CFR S-whz %	10	10	5	2.01	0.94	10.09		
	CFR S-Both %	15	20	15	4.96	1.35	26.77		
Country	% Caseload								
	muac	whz	both						
Senegal	6	88	7	13	18	26	17	13	22
Philippines	12	82	6	16	23	34	20	17	28
Gambia	8	81	10	19	26	37	26	20	32
Kenya	13	78	9	20	28	40	25	21	34
Togo	30	64	6	27	40	55	29	31	46
Mauritania	27	61	11	33	45	59	38	36	51
Mali	24	62	15	35	46	60	42	37	53
India	16	62	21	39	49	61	49	40	55
Benin	26	59	15	38	49	62	44	40	55
Sudan	20	60	20	40	50	63	49	41	57
Burkina Faso	23	58	18	40	51	63	48	42	57
Cameroon	26	57	17	40	51	64	47	42	57
Somalia	31	55	14	40	52	66	46	43	58
South Sudan	27	55	18	42	53	66	50	44	60
Chad	26	51	22	48	58	70	56	49	64
Eritrea	45	45	10	45	59	72	48	52	67
Bangladesh	44	45	11	46	59	73	49	50	65
Nigeria	32	48	19	48	59	72	55	56	70
Niger	26	50	25	50	60	72	59	59	73
Ethiopia	41	43	15	50	63	75	56	54	68
Nepal	37	43	21	53	65	76	60	50	65
Myanmar	37	40	23	57	68	78	64	66	78
Zimbabwe	47	37	16	56	68	79	61	62	76
Liberia	47	36	17	58	70	80	63	49	65
Pakistan	45	35	20	60	71	81	66	61	75
DRC	58	31	11	60	73	83	62	60	74
Tanzania	45	33	23	63	73	83	69	64	77
Guinea	57	28	15	64	75	85	68	68	80
CAR	59	27	13	64	76	85	67	68	80
Haiti	63	26	11	65	77	86	67	71	82
Afghanistan	53	27	20	68	78	86	72	69	81
Angola	56	26	17	68	78	86	71	71	82
Madagascar	54	27	20	68	78	86	73	71	82
Uganda	65	23	12	68	79	88	70	73	84
Burundi	59	24	16	69	79	87	72	72	83
Rwanda	52	25	23	71	80	87	76	72	83
Ivory Coast	63	23	14	70	80	88	72	74	85
Sierra Leone	51	24	25	72	81	88	77	73	84
Zambia	76	17	7	74	84	91	74	79	87
Tajikistan	68	18	14	75	84	91	77	78	87
Malawi	75	16	9	75	85	91	76	79	88
Mozambique	53	17	30	81	87	92	85	82	90

alone and WHZ-alone have about the same mortality risk and that children with both deficits have approximately double the risk. The risks – low MUAC, low WHZ and oedema - appear to be additive; they are not proxies for the same defect. Children with a WHZ <-3Z cannot be described as healthy or less at risk of death than children with a MUAC <115mm.

### Effect of caseload

If there are 100 children with SAM by WHZ with a CFR of 10 per cent and 50 children with SAM by MUAC with a CFR of 20 per cent then, even though the CFR of the MUAC children is double the WHZ CFR, children from each group will have ten SAM-related deaths. As the relative caseload varies widely from country to country, it is a more important determinant of the number of SAM-related child deaths than the relative CFR. We have taken the relative caseloads from Grellety and Golden (2016) and estimated the proportion of all SAM deaths that would be identified and admitted to a MUAC-only programme. We examine the effect of various estimates of CFR for children with a WHZ <-3Z and MUAC <115mm, using 1) theoretical consideration where the CFR is half, the same or double that of the alternative diagnosis and those with both deficits having the sum of the CFRs; and 2) the CFRs derived from the empirical data and the published reports. The results are shown in Table 4. The most likely scenario is given in theoretical simulation B, where the two defects have the same CFR and those with both defects have double the risk of death.

In the countries where most of the children are identified as SAM using WHZ rather than MUAC, most estimates indicate that fewer than half of all SAM-related deaths will be identified in children admitted using a MUAC-only programme, and in most countries only 75 per cent of deaths will be potentially averted if WHZ ceases to be used to identify SAM children.

### Discussion

Our analysis shows that children with SAM identified by WHZ <-3Z and admitted for treatment are at high risk of death; at least as high as those with a low MUAC. In our opinion, the most pressing research needed is to develop simple methods to identify children with a low WHZ in the community, so that these children can be screened and treated. Some innovative methods, based upon photographic technology, are on the horizon and need to be properly funded. At the moment, these children are being neglected and do not feature in MUAC-related statistics such as 'coverage' surveys. To deny that these children are in need of treatment is unethical and in many countries, MUAC-only programmes should not be implemented.

How have we got to this position? It appears to be due to several factors. First, failure to appreciate the effects of mathematical coupling and other confounders that are often severe enough to create a reversal of significance (Tu et al, 2008) – so called Simpson's, Lord's and reversal paradoxes – that result in erroneous conclusions. Second, an exclusive focus on the relative risk of death without consideration of caseload; the relative CFR is not as important as the absolute or relative numbers of SAM deaths. Third, the strong advocacy for the use of MUAC to maximise coverage of treatment programmes has developed into MUAC-only programmes, despite little evidence on the consequences of not admitting low WHZ in different contexts; and lastly, perhaps, due to confirmation bias (Kahneman, 2011; Haselton et al, 2009).

If we assume that there is about an equal mortality for WHZ and MUAC-diagnosed SAM, and that those with both deficits have twice the mortality, then it is possible to estimate the numbers of SAM-related deaths that would be missed globally

**Table 5** Estimation of the possible global deaths from SAM that would be missed using a MUAC-only programme

All-WHZ deaths	Total deaths	WHZ only	MUAC only	Both criteria	MUAC-only % missed	WHZ-only % missed
500,000	756,000	286,000	256,000	214,000	37.7	33.9
1,000,000	1,513,000	571,000	512,000	429,000	33.7	33.8

All-WHZ is the range of global deaths estimated in the Lancet series based on analysis of the numbers of children with SAM by WHZ <-3Z. To this is added the MUAC-only deaths based on the ratios found in Grellety and Golden (2016), using an equal mortality risk for WHZ <-3Z and MUAC <115mm and twice the mortality risk for children with both deficits.

if MUAC-only programmes were to be implemented universally. The results are shown in Table 5. Such a policy would result in between 300,000 and 600,000 SAM deaths occurring in children each year who have no possibility of being treated. This is a very large number of children and suggests that much more analysis should be undertaken in each context before recommending MUAC-only policies.

This in no way should be construed as an attack on the widespread use of MUAC as an independent diagnostic criterion and its merits in enabling screening and increasing treatment coverage at community level; our review reflects that it does not capture a considerable caseload of children who are at risk in different contexts. It is an absolute research priority to develop simple methods of identifying those children at equally high risk who are currently omitted from MUAC-only programmes.

## References

Archie Jr JP. *Mathematic coupling of data: a common source of error*. *Annals of surgery* 1981; 193(3): 296.

Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, Onis M. *Maternal and child undernutrition and overweight in low-income and middle-income countries*. *Lancet* 2013; 382.

Briend A, Alvarez J L, Avril N, Bahwere P, Bailey J, Berkley J A, Binns P, Blackwell N, Dale N, Deconinck H. *Low mid-upper arm circumference identifies children with a high risk of death who should be the priority target for treatment*. *BMC Nutrition* 2016; 2(1): 63.

en-net 2015a. *Only MUAC for admission and discharge?* Emergency Nutrition Network. [www.en-net.org/question/1915.aspx](http://www.en-net.org/question/1915.aspx)

en-net 2015b. *WFH versus MUAC*. Emergency Nutrition Network. [www.en-net.org/question/1922.aspx](http://www.en-net.org/question/1922.aspx)

Grellety E, Golden MH. *Weight-for-height and mid-upper-arm circumference should be used independently to diagnose acute malnutrition: policy implications*. *BMC Nutrition* 2016; 2(1): 10.

Haselton MG, Bryant GA, Wilke A, Frederick DA, Galperin A, Frankenhuys WE, Moore T. *Adaptive rationality: An evolutionary perspective on cognitive bias*. *Social Cognition* 2009; 27(5): 733-763.

Kahneman D. *Thinking, fast and slow*. Macmillan, 2011.

Tu YK, Gunnell D, Gilthorpe MS. *Simpson's paradox, Lord's paradox and Suppression Effects are the same phenomenon – the reversal paradox*. *Emerging Themes in Epidemiology* 2008; 5(1): 2.

Tu YK, Maddick IH, Griffiths GS, Gilthorpe MS. *Mathematical coupling can undermine the statistical assessment of clinical research: illustration from the treatment of guided tissue regeneration*. *Journal of Dentistry* 32(2): 133-142.

WHO (2006). *WHO child growth standards and the identification of severe acute malnutrition in infants and children A Joint Statement by the World Health Organization and the United Nations Children's Fund*.

*This is an extended abstract of three papers under peer review, where the details are described in full. When these papers are published they will be again highlighted in Field Exchange.*

# Postscript

In reviewing this article, the ENN editors posed several questions to Mike and Emmanuel to help our interpretation and understanding of the analysis. Both have kindly agreed to share their feedback. Our questions related to: representativeness of the empirical data analysed regarding extrapolation of risks of SAM-associated deaths identified by different indicators in the community at large; historical evolution of programme effectiveness with improvements in programming over the period during which the data were collected; and impact of admission criteria on the profile of children captured in the dataset (eds).

We have found an ascertainment bias in all the patient studies, including our own, and most of the community cohorts. This is demonstrated by the relative number of children with both MUAC and WHZ deficits (which we call SAM-both) in the subjects analysed and in representative community surveys. Children with SAM-both are more severely malnourished than most of those in the community, but often dominate patient cohorts; hence in-patient and community cohorts are different. We separated SAM-muac and SAM-whz from SAM-both to ameliorate or remove this bias. By eliminating the SAM-both children from the comparison, it is likely that SAM-muac and SAM-whz are much more representative of the situation in the community than would have been the case if we had simply taken all SAM cases together. Because of this we had to eliminate the majority of the children to have a fair comparison of the relative mortality rates. This also removes any bias from the mix of admissions to each facility (i.e. if they were mainly diagnosed by MUAC or WHZ) and makes this consideration irrelevant to the analysis and results. Oedematous children also have to be analysed as a separate group as this is a major confounder.

It is often tacitly assumed that SAM in the community is a fixed reference with which the patients should then be compared, but this is not the case. SAM in the community changes quite markedly with season, food security, economy, violence, epidemics, etc. If the whole community is deteriorating, then the SAM cases that are admitted will be in a worse state and vice versa.

The question we asked is not whether the children directly reflect SAM in the community, but whether children with different degrees of severity of SAM who are diagnosed with either SAM-muac or SAM-whz have a different mortality risk. We were not attempting to compare SAM children with non-SAM children, as many studies have done, but only to compare SAM-muac with SAM-whz; these are quite different questions and require a different study design. Each child with SAM-both could, of course, be counted as both SAM-muac and as SAM-whz – to get a fair comparison of the difference in mortality between the two criteria, they cannot be compared with themselves and appear in both groups being compared!

There are other biases inherent in all such studies. To address the co-morbidity bias, we separated the children into those treated in IPFs, OTPs and SFCs on the basis that the severity and co-morbidity would be IPF > OTP > SFC. We then looked to see if the risk of death with SAM-whz vs SAM-muac was different in the groups with different degrees of co-morbidity. Of course, the case fatality rates were much higher in the IPF than SFC (and SFC may be the same as the community), but the risk of death was not different between those with SAM-whz and SAM-muac in the three modes of treatment; if anything, it was higher in the SAM-whz. The children in each individual facility/programme of course got the same treatment – it was not different in the SAM-whz and the SAM-muac children, so that this could not account for any difference in mortality risk. Comparison of the IPF, OTP and SFC children

also addresses any difference due to the ways of identification of the SAM children and any selection bias that this causes. But there remain potential co-morbidity biases remain when illness affects predominantly different age groups; for example, birth weight, congenital abnormality, HIV and TB are other obvious confounders. Again, this is likely to be different in the three modes of treatment, but the extent of any difference by mode of treatment and their effect on the analysis is not known.

The biggest problem is perhaps verification bias. We do not know how many of the defaulting children died. This is a particular problem with OTP and SFC since the reported death rate is always a minimum death rate as absent children could be alive or dead. Defaulting, transfer of sick children, lost records, lost to follow-up, missing variables, measurement errors, etc. affect all the studies – including the community studies – and need to be taken into account when judging the reliability of the data. We looked to see if there was a difference in degree of this lost data between SAM-muac and SAM-whz. There were minor differences that were inconsistent between SAM-muac and SAM-whz, but not in our opinion sufficient to bias the comparison of mortality from children with MUAC <115mm and WHZ <-3Z. Adding the defaulters to our data does not make a difference to the results that SAM-whz has, in our datasets, a higher mortality risk than SAM-muac. When we used various mortality risks (with SAM-muac either more or less than SAM-whz mortality) with community-based ratios of SAM-muac to SAM-whz with SAM-both factored in, we find a large percentage of deaths would occur in children excluded from treatment using a MUAC-only programme, so any bias in our empirical data does not alter this conclusion.

For more information, contact: Mike Golden, email: [mike@pollgorm.net](mailto:mike@pollgorm.net)