



Estimating the relative contribution of parasitic infections and nutrition for anaemia among school-aged children in Kenya: a subnational geostatistical analysis

Journal:	<i>BMJ Open</i>
Manuscript ID:	bmjopen-2012-001936
Article Type:	Research
Date Submitted by the Author:	06-Aug-2012
Complete List of Authors:	Pullan, Rachel; London School of Hygiene & Tropical Medicine, Department of Disease Control Gitonga, Carol; Kenya Medical Research Institute-Wellcome Trust Research Programme, Mwandawiro, Charles; Eastern and Southern Africa Centre of International Parasite Control, Kenya Medical Research Institute (KEMRI), Snow, Robert; Kenyan Medical Research Institute/University of Oxford-Wellcome Trust Research Programme, Malaria Public Health & Epidemiology Group Brooker, Simon; London School of Hygiene and Tropical Medicine,
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Global health
Keywords:	NUTRITION & DIETETICS, Epidemiology < INFECTIOUS DISEASES, Public health < INFECTIOUS DISEASES, Tropical medicine < INFECTIOUS DISEASES

SCHOLARONE™
Manuscripts

1
2 **Estimating the relative contribution of parasitic infections and nutrition for anaemia**
3
4 **among school-aged children in Kenya: a subnational geostatistical analysis**
5
6

7 Rachel L Pullan¹, Carol Gitonga^{1,2}, Charles Mwandawiro³, Robert W Snow^{2,4} and Simon J Brooker^{1,2}
8
9

10 ¹ *London School of Hygiene and Tropical Medicine, London, UK*

11 ² *Kenya Medical Research Institute-Wellcome Trust Research Programme, Nairobi, Kenya;*

12 ³ *Eastern and Southern Africa Centre of International Parasite Control, Kenya Medical Research Institute*
13 *(KEMRI), Nairobi, Kenya*

14 ⁴ *Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK*
15
16

17
18 **Correspondence:** rachel.pullan@lshtm.ac.uk; Faculty of Infectious and Tropical Diseases, London School
19
20 of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT
21
22

Abstract

Objectives: To quantify geographical variation in the relative contribution of parasitic infections, socio-economic factors and malnutrition in the aetiology of anaemia among school children across Kenya, thereby providing a rational basis for the targeting of an integrated school health package.

Design: Nationally representative cross-sectional survey data were collected using standard protocols. For all included children, data were recorded on haemoglobin concentration and common parasitic infections (*Plasmodium falciparum*, hookworm and schistosomiasis) and socio-economic indicators. Ecological proxies of malnutrition and food security were generated using Demographic and Health Survey and UN Food and Agriculture Organization food security data respectively. Spatially explicit, multilevel models were used to quantify impact upon child haemoglobin concentration.

Setting: Randomly selected schools in ecologically diverse settings across Kenya.

Main outcome measures: Mean haemoglobin concentration adjusted for infection, nutritional and socio-economic risk factors; associated risk ratios and adjusted Population Attributable Fractions (PAFs) for anaemia, by region.

Results: Data were available for 16,941 children in 167 schools; mean Hb was 122.1 g/L and 35.3% of children were anaemic. In multivariate analysis, mean Hb was significantly lower in boys and younger children. Severe malnutrition and interactions between *P. falciparum* and hookworm infections were significantly associated with lower Hb, with greater impacts seen for co-infected children. The contribution of risk factors to anaemia risk varied by province: in 14 year old girls PAFs ranged between 0-27.6% for *P. falciparum*, 0-29.0% for hookworm and 0-18.4% for severe malnutrition.

Conclusions: The observed geographical heterogeneity in the burden of anaemia attributable to different aetiological factors has important implications for the rational targeting of anti-anaemia interventions that can be included into an integrated school health programme.

Article Summary

Article Focus

- Cross-sectional haematological and parasitological school survey data for the period 2008-2010 in Kenya
- Spatially-explicit, Bayesian multi-level models quantifying impact of infection, nutrition and social risk factors on child haemoglobin concentration and adjusted Population Attributable Fractions for anaemia by province.
- Implications of spatial heterogeneity in underlying haemoglobin status and risk factors for cost-effective geographical targeting of anti-anaemia interventions

Key messages

For school-aged children living in Kenya:

- *P. falciparum*, hookworm and schistosomes are all significantly associated with lower mean Hb, with the largest impact seen for coinfection with hookworm and *P. falciparum*, emphasising the importance of parasitic infections in the aetiology of anaemia in African school-age populations
- Poor food intake is suggested to be an important contributor to anaemia in school-aged children.
- There is pronounced geographical heterogeneity in the burden of anaemia attributable to these identified risk factors underscoring importance of appropriately targeting interventions

Strengths and limitations of the study

- Large and uniquely detailed dataset coupled with novel Bayesian modelling framework that allowed for simultaneous quantification of the mean shift in Hb associated with each risk factor, and sub-national risk of anaemia in school-aged children attributable to risk factors.
- Dependence upon cross-sectional design and ecological data limits inference of causality; incorporation of individual, nutritional-related factors and iron status indicators would improve quantification of their involvement.

Background

Children of school-going age (5-14 years) make up one quarter of the population of sub-Saharan Africa,⁽¹⁾ and the proportion of these children attending school has never been higher, in large part due to recent global and national commitments to universal primary education school enrolment.⁽²⁾ Yet many children enrolled in school suffer from poor health and malnutrition and as a result are less likely to attend and complete school or achieve their full educational potential.⁽³⁾ For example, up to 40% of 5-14 year olds are estimated to be anaemic,^(4, 5) which can affect cognition and learning.^(6, 7) Anaemia also affects the health of older girls as they reach puberty, potentially impacting upon future motherhood and the nutrition of the next generation of learners.^(8, 9) Fortunately, a number of simple and cost-effective interventions are available to avert the burden of anaemia, including malaria control,^(10, 11) deworming,⁽¹²⁾ iron supplementation^(13, 14) and school-feeding⁽¹⁵⁾ – interventions which are increasingly delivered as part of an integrated school health package.⁽¹⁶⁾

Identifying the optimal combination of school health interventions requires reliable information on both the geographical distribution of anaemia and the relative contribution of different aetiological factors. A previous study of anaemia among preschool children in West Africa employed a geostatistical framework to interpolate estimates of anaemia among children aged 1-4 years collected through demographic health surveys (DHS), whilst incorporating ecological predictions of malaria and helminth infection. Whilst this work helped define the geographical distribution and limits of anaemia in the region it did not quantify spatial variation in the risk attributable to different aetiological factors.⁽¹⁷⁾ Here we take advantage of uniquely detailed haematological data collected during a comprehensive, national school health survey conducted across Kenya between 2008 and 2010.⁽¹⁸⁾ An initial analysis showed that the relative contribution of *Plasmodium falciparum* varied according to malaria transmission zones.⁽¹⁹⁾ We extend this to investigate the relationships between haemoglobin concentration (Hb) and common parasitic infections (including *P. falciparum*, hookworm and schistosomes), socio-economic indicators and proxies of malnutrition and food security for school children across Kenya. We use an innovative Bayesian approach that quantifies the mean shift in Hb associated with each risk factor, which we use to estimate sub-

1 national variation in the risk of anaemia in school-aged children attributable to malnutrition, malaria and
2
3
4 helminths to define targets for integrated packages of school-health interventions.
5
6
7

8 **Methods**

9 ***School survey data:***

10
11
12 This study uses cross-sectional school survey data collected between September 2008 and March 2010
13
14 using standard protocols, described elsewhere.^(18, 20) In brief, 480 government primary schools were
15
16 selected to provide a national assessment of malaria and anaemia among school children. In a random
17
18 subset of 167 schools stool and urine samples were also collected to assess prevalence and intensity of
19
20 helminth infections. At each school, 10 children plus 1 reserve of each sex were randomly selected from
21
22 each of classes 2-6 from the children present that day. Selected children were asked to provide a
23
24 fingerprick blood sample, which was used to assess haemoglobin (Hb) using a portable photometer
25
26 (Haemocue, Angelholm, Sweden) and *Plasmodium* infection in the peripheral blood based on microscopy-
27
28 corrected malaria rapid diagnostic test (RDT)⁽¹⁸⁾ results. Children in all schools were asked to provide stool
29
30 samples which were examined for the eggs of intestinal nematodes (*Ascaris lumbricoides*, *Trichuris*
31
32 *trichiura* and hookworm species) and *S. mansoni* using the Kato-Katz method. In schools on the coast,
33
34 where *Schistosoma haematobium* is widespread, children were asked to provide urine samples which
35
36 were examined by urine filtration. A questionnaire was administered to all pupils to obtain data on
37
38 insecticide-treated net ownership and use, history of anthelmintic medication, key socioeconomic
39
40 variables, household construction and education of the child's parent/guardian. School locations were
41
42 determined using a hand-held Global Positioning System and classified as urban using an updated 2010
43
44 urban extents (UE) mask derived from GRUMP.^(21, 22)
45
46
47
48
49
50

51 ***Nutritional proxy data***

52
53
54
55 Anthropometric and food intake data were not available for all of our study population, and we instead
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

1 derived from the Kenya 2008-09 Demographic and Health Survey (DHS).^(23, 24) Anthropometric
2 measurements were extracted for children ≤ 5 years and for women of child-bearing age (ages 15 to 49)
3 (see Table 1). For children ≤ 5 years, the reported proportion consuming food rich in iron (flesh and organ
4 meat) in the past 24 hours was also extracted. These indicators, which provide a general measure of
5 nutritional status in the wider community, were geo-located to the centre of each sampling cluster.
6
7 Bayesian kriging⁽²⁵⁾ was used to interpolate a distribution of possible values for each indicator to survey
8 school locations using the Bayesian statistical software WinBUGS version 14.1 (Medical Research Council
9 Biostatistics Unit and Imperial College London). Mean predicted prevalence values were then categorised
10 according to WHO-defined cut-off values for severity of malnutrition by prevalence range (Table 1). See
11 Technical Appendix for further details.
12
13
14
15
16
17
18
19
20
21
22
23
24
25

26 Two additional indicators were used to describe general food availability at each school location:
27 predominant source of livelihood (based on 2010 FAO Livelihood zones^(26, 27)); and food security
28 conditions at the time of the survey (based on the FAO Food Insecurity Severity Scale^(27, 28)). Both
29 indicators were obtained from the Famine Early Warning System Network (FEWS NET;
30 <http://www.fews.net>). Information on the intended provision of formal, government-run school-feeding
31 programmes was abstracted from the Homegrown School Meals Programme (HGSMP) and the Emergency
32 HGSMP 2010 registers provided by the Kenya Ministry of Education (*personal communication*).
33
34
35
36
37
38
39
40
41
42
43

44 **Analysis**

45
46 A schematic summarising the major steps in our analysis strategy is provided in Figure 1. Haemoglobin
47 levels were adjusted for altitude and individuals were classified as anaemic using age and sex specific cut-
48 offs.⁽²⁹⁾ Children were classified as having medium/high intensity hookworm infection if infection intensity
49 exceeded 100 epg, and medium/high intensity *S. haematobium* infection if infection intensity exceeded
50 500 epg (both representing the top 10th percentile). Reported information on ownership of household
51
52
53
54
55
56
57
58
59
60

1 assets and household construction was used to generate wealth indices using principle component
2 analysis,⁽³⁰⁾ and resulting scores were divided into quintiles.
3
4
5
6
7

8 The initial candidate set of predictor variables for univariate analysis included age, sex, relative socio-
9 economic status, education-level of the household head, protective behaviours (bednet usage and
10 deworming) and infection status with hookworm, schistosome and *P. falciparum*, in addition to school-
11 level ecological covariates including urban/rural location, predicted prevalence of anthropometric
12 indicators in younger children and women of reproductive age, livelihood zone and food security status
13 and intended provision of school-feeding. In order to select candidate variables for spatial multivariable
14 analysis, univariate analysis was undertaken using logistic regression (criteria: Wald test $P < .1$) using Stata
15 11.1 software (StataCorp). Standard errors were adjusted for dependence between individuals within
16 schools using a clustered sandwich estimator. Backward-stepwise elimination was then used to generate a
17 minimum adequate multivariate model; excluded covariates ($P > .05$) were retested in the minimal model
18 to confirm lack of association.
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

35 Retained covariates were included in a spatial multivariable Bayesian mixed-effects linear regression
36 model using WinBUGs 14.1. This model included individual and school-level ecological fixed effects and a
37 school-level geostatistical random effect (using an isotropic, stationary exponential decay function),⁽²⁵⁾
38 thereby incorporating a degree of spatial smoothing. Uncertainty in the predicted prevalence of
39 anthropometric proxy indicators was incorporated by modelling these covariates as random variables
40 with a beta distribution, defined using the prediction posterior distribution. At each model realisation,
41 malnutrition was categorised as either low, medium/high or severe for each school location, according to
42 Table 1.
43
44
45
46
47
48
49
50
51
52
53
54

55 Given the large observed differences in mean Hb and prevalence of predictors by region and demographic
56 group, adjusted PAFs for anaemia were estimated for two indicator demographic groups (boys aged 7
57
58
59
60

1 years, and girls aged 14 years) for each province in Kenya. In brief, at each model realisation we estimated
2
3 the prevalence of anaemia in each of the risk groups based upon the posterior mean and standard
4
5 deviation of the Hb distribution within each school. These were used to estimate (i) the relative risk of
6
7 anaemia for each risk group compared with baseline risk, and (ii) the associated adjusted PAF, as
8
9 described by Rockhill *et al* (1998).⁽³¹⁾ Estimates were subsequently summarised by province. Full
10
11 statistical notation of Bayesian geostatistical modelling, PAF and model validation procedures are
12
13 presented in Technical Appendix. Model validation using hold out datasets, presented in detail in the
14
15 accompanying Technical Appendix, suggest very high degree of correlation at the school level between
16
17 observed and expected mean Hb levels ($R^2 > 0.98$), and prevalence of anaemia ($R^2 > 0.95$) suggesting the
18
19 approach we used to estimate anaemia prevalence (and associated attributable fractions) was
20
21 appropriate.
22
23
24
25
26
27
28
29

30 Results:

31
32 Data on Hb, malaria parasitaemia and helminth infection were available for 16,941 children (aged 5-16) in
33
34 167 schools (Figure 2). Mean altitude-adjusted Hb concentration was 122.1 g/L, with 35.3% of children
35
36 estimated to be anaemic. The spatial distribution of the unadjusted Hb data is presented in Figure 2,
37
38 summarised by school. There were significant geographical heterogeneity in mean Hb across the country
39
40 (Moran's $I p < 0.001$), with the lowest average levels observed in Coast Province (120.1 g/L) and the highest
41
42 in Central Province (134.0 g/L). Univariate analysis suggested that Hb levels increased with age and with
43
44 relative socio-economic status, and were significantly lower in individuals infected with *P. falciparum* or
45
46 medium/high intensity helminth infections (Table 1). Hb was also significantly lower in areas with
47
48 predicted severe malnutrition, although there was no ecological relationship with predicted consumption
49
50 of iron rich foods. Univariate analysis also suggests a relationship between Hb and food availability, with
51
52 significantly lower Hb levels in children living in very food insecure areas.
53
54
55
56
57
58
59
60

Bayesian hierarchical model of haemoglobin

Table 3 presents a multivariate analysis of factors associated with Hb levels. On average, boys had lower mean Hb compared with girls by 1.53 g/L, and older children tended to have higher Hb levels, increasing 1.08 g/L each yearly age interval. There was strong evidence for a multiplicative negative interaction between infection status and Hb levels, highlighting the health impacts of parasitic infections in school-aged children. Children with medium/high intensity hookworm infection had Hb levels 1.03 g/L lower than uninfected children, those with malaria 3.00 g/L lower and *S. haematobium* infection 1.75 g/L; children co-infected with both malaria and medium/high intensity hookworm infection had Hb levels 5.56 g/L lower than uninfected children.

Proxy measures of nutritional status were also strongly associated with Hb levels. School-children living in areas predicted to have severe prevalence of wasting had Hb levels that were on average 9.74 g/L lower than those living in low risk areas. School-feeding programmes also appear to have been targeted to districts with increased risk of anaemia, although examination of the 95% credible intervals suggests no statistical difference within target districts between mean Hb levels in schools with and without intended provision of formal school-feeding programmes. Perhaps surprisingly, food security at the time of the survey (a qualitative measure that allows adjustment for acute conditions and the impact of seasonality) was not associated with Hb levels in multivariate analysis. Similarly, after accounting for all other factors in the model, relative socio-economic status was not significantly associated with Hb levels.

The random effects variance components from this model indicate that most of the residual variation remains between children within schools: of the total unexplained variability in Hb, 13.8% occurred at the school level, whilst 86.2% occurred between children within schools. Closer examination of the spatial school-level random effect suggests that, after accounting for included predictors, most schools (81%) had similar Hb levels that could not be distinguished statistically (shown in green in Figure 2). There was however some residual spatial clustering of schools with higher/lower than expected Hb levels (foci radii

1 of ~190 km, Table 2) with significantly lower mean Hb levels observed in schools in the far south of Coast
2 Province and in Nyanza and Western provinces. Interestingly, comparison of observed prevalence of *P.*
3 *falciparum* and medium/high intensity hookworm suggests that prevalence of both infections was higher
4 for these schools than for other schools in these provinces (*P. falciparum* Kruskal-Wallis p 0.02;
5 medium/high intensity hookworm Kruskal-Wallis p 0.01). In contrast, mean Hb levels were significantly
6 higher for schools in Rift Valley, Central and Eastern provinces
7
8
9
10
11
12
13

14 **Adjusted Population Attributable Risk Fractions**

15
16
17 Using the final Bayesian hierarchical model, we estimated for two demographic groups (girls aged 14
18 years, and boys aged 7 years) the PAF of anaemia due to *P. falciparum*, medium/high intensity hookworm
19 infection, medium/high intensity *S. haematobium* infection, and acute malnutrition (prevalence of
20 wasting in under 5s exceeding 15%) for each province in Kenya, adjusting for all other included predictors
21 (Table 4). Our results suggest that the relative contribution of each risk factor varies considerably by
22 province. For example, in Coast Province 13% (95% BCI: 4-23%) of anaemia in 14 year old girls attending
23 school may be attributable to medium/high intensity hookworm, comparing with 6% (95% BCI: 2-10%) to
24 *P. falciparum* infection and 3.1% (95% BCI: 0-8%) to schistosomiasis. Meanwhile, results suggest that
25 reduction in acute malnutrition in Coast Province (to <5% wasting in \leq 5s) would reduce anaemia
26 prevalence by 19% (95% BCI: 0-38%). This contrasts with Western Province, where the estimated risk of
27 anaemia attributable to medium and high intensity hookworm is 29% (95% BCI: 1-55%) and to *P.*
28 *falciparum* 24% (95% BCI 0-41%); no school children in this district are exposed to *S. haematobium* and
29 acute malnutrition in under 5s is predicted not to exceed 5%. The wide 95% BCIs observed reflect (i) the
30 smaller sample sizes for provinces other than Coast (given that N = schools, not individuals), and (ii) the
31 additional uncertainty surrounding proxies for malnutrition. Trends in PAFs observed for predictors in 7
32 year old boys mirror those for older girls but are generally lower, a reflection of the lower baseline Hb
33 levels in this group.
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discussion

As global commitment and resources for school health and nutrition initiatives grow,⁽¹⁶⁾ there is an urgent need to strengthen our understanding of the likely impact of interventions aimed at tackling important health and development outcomes such as anaemia. This requires a detailed understanding of the relative contribution of aetiological factors at policy relevant scales. Here we present a uniquely detailed national analysis of risk factors for anaemia among schoolchildren in 167 schools across Kenya. We introduce a new approach using multilevel linear modelling to provide a detailed description the varying contribution of risk factors by administrative unit (province), whilst fully accounting for spatial variation in underlying mean Hb. Results provide a vital decision-making tool for cost-effective targeting the delivery of integrated, packages of interventions tailored to each province, with the aim of reducing anaemia in schoolchildren.

Our analysis found evidence that *P. falciparum*, hookworm and schistosomes were each significantly associated with lower mean Hb, observations consistent with current knowledge. Such findings corroborate recent results from the Global Burden of Disease study, which suggests that globally for anaemia 10.2% of years lived with a disability (YLD) across all ages could be attributed to hookworm, 17.9% to malaria and 3.4% to schistosomiasis.⁽³²⁾ Our analysis provides three additional insights into the epidemiology of anaemia. First, we confirm small-scale studies⁽³³⁻³⁷⁾ that have suggested that coinfection with hookworm and *P. falciparum* may have a multiplicative effect on Hb level – effects which occur across Kenya. Second, we demonstrate pronounced geographical heterogeneity in the burden of anaemia attributable to identified risk factors. Associations for hookworm and *P. falciparum* were most pronounced in Western Province, where results suggest that together these two infections were attributable for close to 53% of anaemia cases in girls aged 14 years. Together, these findings emphasise the importance of parasitic infections in the aetiology of anaemia in school-aged children and underscore current efforts to control helminths and malaria as part of integrated school health programmes.^(38, 39) The

1 PAFs provided here are useful for translating our findings into numbers that can help policy decision
2
3 makers appreciate the potential benefits of carefully targeted programmes tailored to each province that
4
5 strike the best balance between dietary interventions, deworming and malaria control.^(17, 40)
6
7
8
9

10 Third, whilst acknowledging that we rely on proxy indicators, our results do suggest that poor food intake
11
12 is an important contributor to anaemia in school children: the final model indicates that in communities
13
14 where acute malnutrition is a severe public health problem (i.e. where wasting in under 5s exceeds 15%),
15
16 mean Hb levels in school-aged children are on average 9.7 g/L lower than well nourished communities,
17
18 and as such eliminating acute malnutrition as a public health problem could avert close to 20% of anaemia
19
20 cases in 14 year old girls. This is consistent with limited evidence from smaller individual-level studies
21
22 suggesting an association between under-nutrition and anaemia in school-aged populations.^(41, 42) Our
23
24 results also provide some indication that whilst existing school-feeding programmes in Kenya do appear
25
26 to be targeted to those at most need, they are currently ineffective at tackling anaemia in this population.
27
28
29
30
31
32

33 Our spatially explicit hierarchical approach seeks to maximise the usefulness of this observational, cross-
34
35 sectional data and results show that whilst much of the between school variation could be explained by
36
37 included risk factors, a high degree of variation does still remains within schools. This suggests that there
38
39 are individual-level factors beyond those measured that still need to be considered. In addition to
40
41 individual iron status or food intake there are a number of candidates. HIV-infection is potential likely to
42
43 be a contributor, as even in asymptomatic HIV infected individuals anaemia is a common manifestation.⁽⁴³⁾
44
45 ⁴⁴⁾ We were also unable to account for genetic traits such as sickle-cell and other haemoglobinopathies,⁽⁴⁵⁾
46
47 or for the potential effect of menarche in adolescent girls.⁽⁴⁶⁾
48
49
50
51
52

53 A number of additional study limitations should also be acknowledged. As we have previously
54
55 acknowledged,⁽²⁰⁾ cross-sectional studies such as this are inevitably subject to a number of limitations
56
57 including the potential of ecological fallacy, especially when incorporating population-level estimates. We
58
59
60

1 acknowledge that our indicator of school-feeding (inclusion in the government school-feeding registers) is
2 far from optimal, and the study would have benefitted from detailed individual-level data on iron status
3 and food intake. In addition, modelled associations are likely to be lower than the true impact, due to
4 non-systematic measurement error in our infection indicators, which rely on single samples using
5 standard parasitological diagnosis with <100% sensitivity.^(47, 48) Despite a robust stratified random
6 sampling frame, our results may also be subject to some selection bias, as very ill children such as those
7 with acute infections and severe chronic malnutrition may have been absent on the day of the school
8 survey and thus not included in the study sample. Neither is the analysis able to account for the impact of
9 recently cleared infections on current Hb status. Finally, the current results represent a snapshot of the
10 period 2008-2010 and it is conceivable that seasonality and intervention-related factors may have
11 confounded results. However, the inclusion of mosquito net use, recent deworming or food security (a
12 seasonal covariate) in multivariate models did not improve model fits, suggesting that temporal and
13 control factors were relatively unimportant.

31 Conclusions

32 A recent focus on school-child and adolescent health has emphasised the need to develop a deeper
33 understanding of factors impacting health and development of this population group.⁽⁴⁹⁻⁵¹⁾ Here we
34 present a novel two-step process to model Hb levels and then predict associated anaemia prevalence in
35 different areas of Kenya. Our approach adds considerable value over and above what could be achieved
36 modelling anaemia directly, and the adopted geostatistical Bayesian framework allows for propagation of
37 uncertainty associated at each step (including generation of predicted ecological covariates) into the final
38 confidence limits. Future expansion of this model could look at interpolating existing results to make
39 predictions in non-surveyed areas, although this would require generation of values for all included
40 individual-level covariates for each prediction location, increasing prediction uncertainty. Our results
41 highlight the marked geographical heterogeneity in the burden of anaemia attributable to identified risk
42 factors in Kenya, which has major implications for effective targeting of packages of school-based
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1 interventions. The priority now should be to explore in detail the health and educational impact of
2
3 carefully designed integrated infection and nutritional interventions in this age group.
4
5
6
7

8 **Acknowledgements**

9
10 We are very grateful to all the teachers, parents and especially children who participated in this survey
11
12 and to the many facilitators and members of the survey teams who collected the data. The survey work
13
14 was supported by the Wellcome Trust UK as part of fellowship support to SJB (#81673) and by the Division
15
16 of Malaria Control, Ministry of Public Health and Sanitation through a grant from the Department for
17
18 International Development (DfID) through the WHO Kenya Country Office to Dr AM Noor and Dr E Juma.
19
20
21
22

23 **Footnotes**

24
25 **Contributors** RLP was responsible for designing and conducting analysis, interpretation, drafting and
26
27 production of the final manuscript. CG contributed to the survey design, data assembly and cleaning and
28
29 contributed to final manuscript. CM and RWS contributed to survey design, interpretation of results and
30
31 contributed to final manuscript. SJB provided scientific guidance and contributed to interpretation and
32
33 preparation of the final manuscript. All authors read and approved the final manuscript.
34

35 **Funding** The fieldwork was supported by the Wellcome Trust UK and by the Division of Malaria Control,
36
37 Ministry of Public Health and Sanitation. RLP acknowledges support of the Bill & Melinda Gates
38
39 Foundation and CWG is supported by a Commonwealth Scholarship from the DfID. SJB is supported by a
40
41 Wellcome Trust Senior Fellowship in Basic Biomedical Science (098045) and RWS is supported by a
42
43 Wellcome Trust Principal Fellowship (079080).
44

45 **Competing interests** None.
46
47

48 **Data sharing statement:** The school survey dataset used for this analysis is available for download
49
50 through the Global Atlas of Helminth Infection (www.thiswormyworld.org). The WinBUGs statistical code
51
52 is available from the corresponding author at rachel.pullan@lshtm.ac.uk.
53
54
55
56
57
58
59
60

References

1. United Nations Department of Economic and Social Affairs. World population prospects, 2010 revision; Available from: <http://esa.Un.Org/unpd/wpp/>.
2. Global Partnership for Education. All children learning: Report 2011. Global Partnership for Education; Available from: <http://allchildrenlearningreport.globalpartnership.org>
3. World Bank. Fresh: Focusing resources on education and school health. A fresh start to improving the quality and equity of education. UNESCO, UNICEF, WHO & World Bank; 2000
4. McLean E, Cogswell M, Egli I, Wojdyla D, de Benoist B. Worldwide prevalence of anaemia, WHO vitamin and mineral nutritional information system, 1993-2005. *Public Health Nutrition*. 2008;12(4):444-454.
5. Best C, Neufingerl N, van Geel L, van den Briel T, Osendarp S. The nutritional status of school-aged children: Why should we care? *Food and Nutrition Bulletin*. 2010;31(3):400-417.
6. Grantham-McGregor S, Ani C. A review of studies on the effect of iron deficiency on cognitive development in children. *Journal of Nutrition*. 2001 Feb;131(2S-2):649S-666S.
7. Fretham SJ, Carlson ES, Georgieff MK. The role of iron in learning and memory. *Advances in Nutrition*. 2011;2(2):112-121.
8. Scholl TO. Iron status during pregnancy: Setting the stage for mother and infant. *American Journal of Clinical Nutrition*. 2005;81(5):1218S-1222S.
9. Dharmalingam A, Navaneetham K, Krishnakumar CS. Nutritional status of mothers and low birth weight in india. *Maternal and Child Health Journal*. 2010;14(2):290-298.
10. Clarke SE, Jukes M, Njagi JK, Khasakhala L, Cundill B, Otido J, et al. Effect of intermittent preventive treatment of malaria on health and education in schoolchildren: A cluster-randomised, double-blind, placebo-controlled trial. *Lancet*. 2008;372(9633):127-138.
11. Leenstra T, Phillips-Howard PA, Kariuki SK, Hawley WA, Alaii JA, Rosen DH, et al. Permethrin-treated bed nets in the prevention of malaria and anemia in adolescent schoolgirls in western kenya. *American Journal of Tropical Medicine and Hygiene*. 2003;68(S86-S93).
12. Smith JL, Brooker S. The impact of hookworm infection and deworming on anaemia in non-pregnant populations: A quantitative update. *Tropical Medicine and International Health*. 2010;15(7):776-795.
13. Leenstra T, Kariuki SK, Kurtis JD, Oloo AJ, Kager PA, ter Kuile FO. The effect of weekly iron and vitamin a supplementation on hemoglobin levels and iron status in adolescent schoolgirls in western kenya. *European Journal of Clinical Nutrition*. 2009;63(2):173-182.
14. Gera T, Sachdev HP, Nestel P. Effects of iron supplementation on physical performance in children and adolescents: Systematic review of randomized controlled trials. *Indian Pediatrics*. 2007;44:15-24.
15. Bundy DAP, Burbano C, Grosh M, Gelli A, Jukes M, Drake L. *Rethinking school feeding: Social safety nets, child development and the education sector*. The World Bank: New York; 2009.
16. Bundy DAP. *Rethinking school health: A key component of education for all*. Washington: The World Bank; 2012.
17. Soares Magalhães RJ, Clements ACA. Mapping the risk of anaemia in preschool-age children: The contribution of malnutrition, malaria, and helminth infections in West Africa. *PLoS Medicine*. 2011;8(6):e1000438.
18. Gitonga C, Karanja P, Kihara J, Mwanje M, Juma E, Snow R, et al. Implementing school malaria surveys in kenya: Towards a national surveillance system. *Malaria Journal*. 2010;9(1):306.
19. Gitonga CW, Edwards T, Karanja PN, Noor AM, Snow RW, Brooker SJ. *Plasmodium* infection, anaemia and mosquito net use among school children across different settings in Kenya. *Tropical Medicine and International Health*. 2012;17(7):858-870.
20. Brooker S, Pullan RL, Gitonga CW, Ashton RA, Kolaczinski JH, Kabatereine NB, et al. *Plasmodium*-helminth coinfection and its sources of heterogeneity across east africa. *Journal of Infectious Diseases* 2012;205(5):841-852.
21. Balk DL, Deichmann U, Yetman G, Pozzi F, Hay SI, Nelson A. Determining global population distribution: Methods, applications and data. *Advances in Parasitology*. 2006;62:119-156.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
22. Balk D, Pozzi F, Yetman G, Deichmann U, Nelson A. The distribution of people and the dimension of place: Methodologies to improve the global estimation of urban extents. Urban Remote Sensing Conference; 2004; Available at: <http://www.iussp.org/Activities/wgc-urb/balk.pdf>
 23. Kenya National Bureau of Statistics, Measure DHS. Kenya: Dsh, 2008-09 - final report. Calverton, Maryland, USA: Measure DHS, ICF Macro; 2010
 24. Measure DHS. Kenya 2008-09 DHS data. Available from: <http://www.measuredhs.com>
 25. Diggle P, Moyeed R, Tawn J. Model-based geostatistics (with discussion). *Applied Statistics*. 1998;47:299-350.
 26. FEWSNET. Famine early warning system network (FEWSNET) livelihood zone map for Kenya. FEWSNET and USAID; 2011
 27. Grillo J, Holt J. Application of the livelihood zone maps and profiles for food security analysis and early warning: USAID FEWSNET; 2009
 28. FEWSNET. Famine early warning system network (FEWSNET) Kenya archive: Food security update. FEWSNET and USAID; 2011
 29. WHO. Iron deficiency anaemia: Assessment, prevention and control. A guide for programme managers. Geneva: World Health Organisation; 2001
 30. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data--or tears: An application to educational enrollments in states of india. *Demography*. 2001;38(1):115-32.
 31. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *American Journal of Public Health*. 1998;88(1):15-19.
 32. Jasrasaria J, Murray CJL, Naghavi M. Global Burden of Disease 2010 Estimate: Strategy Report for Anaemia Envelope. [Unpublished report]. *Personal Communication* 2012.
 33. Brooker S, Akhwale WS, Pullan R, Estambale B, Clarke S, Snow RW, et al. Epidemiology of *Plasmodium*-helminth coinfection in Africa: Potential impact on anaemia and prospects for combining control. *American Journal of Tropical Medicine & Hygiene*. 2007;77:88-98.
 34. Pullan R, Brooker S. The health impact of polyparasitism in humans: Are we under-estimating the burden of parasitic diseases? *Parasitology*. 2008:1-12.
 35. Nkuo-Akenji TK, Chi PC, Cho JF, Ndamukong KKJ, Sumbele I. Malaria and helminth co-infection in children living in a malaria endemic setting of Mount Cameroon and predictors of anaemia. *Journal of Parasitology*. 2006;92(6):1191-1195.
 36. Midzi N, Mtapuri-Zinyowera S, Mappingure MP, Sangweme D, Chirehwa MT, Brouwer KC, et al. Consequences of polyparasitism on anaemia among primary school children in Zimbabwe. *Acta Trop*. 2010;115:103-111.
 37. Ezeamama EE, McGarvey ST, Acosta LP, Zierler S, Manalo DL, Wu H-W, et al. The synergistic effect of concomitant schistosomiasis, hookworm, and trichuris infections on children's anemia burden. *PLoS Negl Trop Dis*. 2008;2(6):e245.
 38. Bundy DAP, Shaeffer S, Jukes M, Beegle K, Gillespie A, Drake L, et al. School based health and nutrition programs. *Chapter 58 in "Disease Control Priorities for Developing Countries"*. 2006.
 39. Brooker S, Clarke S, Snow RW, Bundy DAP. Malaria in african schoolchildren: Options for control. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2008;102(4):304-305.
 40. Rowe AK, Powell KE, Flanders WD. Why population attributable fractions can sum to more than one. *American Journal of Preventative Medicine*. 2004;26:243-249.
 41. Anumudu C, Adolami M, Igwe C, Nwagwu M, Keshinro O. Nutritional anaemia and malaria in pre-school and school-aged children. 2008;1(11-17).
 42. Daboné C, Delise HF, Receveur O. Poor nutritional status of schoolchildren in urban and peri-urban areas of Ouagadougou (Burkina Faso). *Nutrition Journal*. 2011;10:34.
 43. Calis JC, van Hensbroek MB, de Haan RJ, Moons P, Brabin BJ, Bates I. HIV-associated anemia in children: A systematic review from a global perspective. *AIDS*. 2008;22(10):1099-1112.
 44. Levine AM, Berhane K, Masri-Lavine L, Sanchez M, Young M, Augenbraun M, et al. Prevalence and correlates of anemia in a large cohort of HIV-infected women: Women's interagency HIV study. *Journal of Acquired Immune Deficiency Syndrome*. 2001;26:28-35.
 45. Tolentino K, Friedman JF. An update on anaemia in less developed countries. *American Journal of Tropical Medicine and Hygiene*. 2007;77:44-51.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
46. Leenstra T, Kakiuki SK, Kurtis JD, Oloo AJ, Kager PA, ter Kuile FO. Prevalence and severity of anaemia and iron deficiency: Cross-sectional studies in adolescent schoolgirls in western Kenya. *European Journal of Clinical Nutrition*. 2004;58(4):681-691.
47. Booth M, Vounatsou P, N'Goran E K, Tanner M, Utzinger J. The influence of sampling effort and the performance of the kato-katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Cote d'Ivoire. *Parasitology*. 2003;127(Pt 6):525-31.
48. Okell LC, Ghani AC, Lyons E, Drakeley C. Submicroscopic infection in *Plasmodium falciparum* endemic populations: A systematic review and meta-analysis. *Journal of Infectious Disease*. 2009;15(10):1509-1517.
49. World Bank. World development report 2007: Development and the next generation. Washington: The World Bank
50. Little AW, Green A. Successful globalisation, education and sustainable development. *International Journal of Education and Development*. 2009;29:166-174.
51. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S-J, Dick B, Ezech AC, et al. Adolescence: A foundation for future health. *The Lancet*. 2012;379(9826):1630-1640.
52. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiological research. Belmont, CA: Lifetime Learning Publications; 1982. p. 163.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure legends:

Figure 1 Schematic of adopted analysis strategy. The dual aim of this analysis strategy was to (i) model the impact of risk factors on children's haemoglobin (Hb) levels, and (ii) estimate the adjusted Population Attributable Risk Fraction (PAF) associated with important, modifiable risk factors. This process was repeated at each realisation of the Bayesian hierarchical model to provide posterior mean estimates and credible intervals for all parameters of relevance.

Figure 2 Spatial distribution of mean haemoglobin distribution across Kenya, by school. HB is adjusted for altitude, population density (persons / km²) is shown for reference.

Figure 3 Graph of spatial random-effect (school-level residual) from the multivariate, Bayesian hierarchical model for Hb. Each point represents one school; the posterior mean is shown by the dot and the 95% BCI are represented by the error bars. The horizontal straight line in the middle of the graph represents the overall mean of the residuals, which are zero centred. Inset shows the school locations, shaded by the value of the school-level residual.

For peer review only

Tables:

Table 1 Definition of included anthropometric indicators, and population-based cut-off values for severity of malnutrition, as defined by the World Health Organization.

Anthropometric Measurement		Severity of malnutrition (prevalence range for < 6 years)		
		Low severity	Medium/high severity	Very high severity
Stunting	Height-for-age Z-score > -2 SD	<20%	20-40%	>40%
Wasting	weight-for-height Z-score > -2 SD	<5%	5-15%	>15%
Underweight	weight-for-age <i>or</i> BMI-for-age Z-score > -2 SD	<10%	10-30%	>30%

Table 2 Mean Hb concentration according to infection status, nutritional proxies and other characteristics in 16,941 children aged 5-16 years (univariate analysis).

Variable	Data type and source	Children No (%)	Mean Hb shift, g/L (SE)	p
Sex				
Female	Individual – KHS	8,314 (49.1)	(vs. female)	
Male	Individual – KHS	8,627 (50.9)	- 1.36 (0.28)	<0.001
Age (in years)	Individual – KHS	-	+ 0.89 (0.11)	<0.001
Socio-economic indicators				
Educated household head ¹	Individual – KHS	8,333 (49.2)	+ 0.82 (0.41)	0.05
Household owns a mobile phone	Individual – KHS	10,065 (59.4)	+ 1.42 (0.40)	<0.001
SES quintile				
Lowest	Individual – KHS	3,272 (20.4)	(vs. lowest)	
Mid (quintiles 2-4)		9,926 (61.8)	+ 2.35 (0.50)	<0.001
Highest		2,854 (17.8)	+ 3.07 (0.82)	<0.001
Urban location	School-level – digital UE map	1,237 (7.3)	1.56 (1.52)	0.3
Current infections				
Malaria parasitaemic	Individual – KHS	607 (3.6)	-3.36 (0.77)	<0.001
Med/high intensity hookworm (>100 epg)	Individual – KHS	931 (5.5)	-2.81 (0.69)	<0.001
Med/high intensity <i>S. haematobium</i> (>500 epg)	Individual – KHS	799 (4.7)	-3.81 (1.02)	<0.001
Usually sleeps under a net	Individual – KHS	9,763 (57.6)	- 1.78 (0.52)	0.001
Taken anthelmintic in previous 6 months	Individual – KHS	7,494 (44.2)	+ 1.24 (0.48)	0.01
Nutritional proxies				
Wasting in under 5s (%) ²				
Under 5%	Ecological – interpolated	4,027 (23.8)	(vs <5%)	
5-15%	from 2008 DHS	10,257 (60.6)	-5.58 (1.04)	<0.001
Over 15%		2,657 (15.7)	-8.14 (1.21)	<0.001
Low BMI in adult women (%) ³				
Under 20%	Ecological – interpolated	12,867 (76.0)	(vs. <20%)	
Over 20%	from 2008 DHS	4,072 (24.0)	-1.98 (0.80)	0.01
Consumption iron rich foods in under 5s (%) ⁴				
Under 25%	Ecological – interpolated	14,572 (86.0)	(vs. <25%)	
Over 25%	from 2008 DHS	2,369 (14.0)	+ 1.07 (1.24)	0.4
School-feeding (SF) status: ⁵				
Not a SF district	School-level – MoE registers	3,690 (21.8)	(vs. not SF dist)	
SF district – no programme in school		8,417 (49.7)	-7.34 (1.08)	<0.001
SF district – programme in school		4,834 (28.5)	-4.80 (1.19)	<0.001
Generalised livelihood zone ⁶				
Mixed farming		3,295 (19.5)	(vs. mixed)	
Pastoral and agro-pastoral	Ecological – FEWSNET	4,991 (29.5)	+ 1.11 (1.45)	0.4
Riverine and fishing		641 (3.8)	+ 2.08 (2.42)	0.4
Medium potential farming		4,073 (24.0)	- 2.93 (1.13)	0.02
High potential farming		3,941 (23.3)	+ 1.37 (2.14)	0.2
Food security status ⁷				
Generally food secure	Ecological – FEWSNET	2,756 (16.3)	(vs. secure)	
Moderately food insecure		7,755 (45.8)	-3.20 (1.27)	0.01
Highly food insecure		6,048 (35.7)	- 6.17 (1.25)	<0.001
Extremely food insecure		382 (2.3)	- 5.12 (1.52)	0.001

KSH; Kenya School Health survey (conducted 2008-2010). DHS; Kenya Demographic Health Survey conducted 2008/2009. MoE; Kenya Ministry of Education. FEWSNET; Famine early warning system network. SE; standard error.

¹ household head educated to primary complete or above (reported).

^{2,3,4} Interpolated from Kenya 2008/2009 DHS survey

² school included in either the Homegrown School Meals Programme (HGSM) or Emergency HGSM 2010 registers

⁶ based on the USAID/FEWSNET Generalised Livelihood Zone 2010 Classification for Kenya

⁷ based on the USAID/FEWSNET Food Insecurity Severity Scale for the quarter corresponding to the KSH survey date.

Table 3 Bayesian hierarchical multivariate regression model for haemoglobin levels.

Variable	Mean Hb shift (g/L)	95% BCI / SD
Baseline haemoglobin	117.3	(113.9, 120.5)
Sex (male)	-1.53	(-1.94, -1.13)
Age (in years)	+1.08	(+0.99, +1.17)
SES quintile (vs. quintiles 2-4)		
lowest	-0.56	(-1.14, +0.02)
highest	0.25	(-0.31, +0.86)
Med/high intensity <i>S. haematobium</i> (>500 epg)	-1.75	(-2.92, -0.60)
Med/high intensity hookworm (>100 epg)	-1.03	(-2.00, -0.02)
Malaria parasitaemic	-3.00	(-4.29, -1.68)
Med/high hookworm – malaria co-infection	-5.56	(-8.64, -2.67)
Wasting in under 5s (vs <5% prevalence)		
5-15%	-4.39	(-6.56, -2.09)
>15%	-9.74	(-12.16, -7.45)
School-feeding (SF) status:		
SF district – no programme in school	-2.14	(-4.76, +0.46)
SF district – programme in school	-2.78	(-5.43, -0.21)
Random effect terms		
Individual σ^2	183.6	(2.02)
Spatial σ^2 (school-level)	29.5	(7.93)
Range of spatial correlation (km)	194	(99,318)

Table 4: Predicted adjusted population attributable risk fractions for key predictors across Kenya. Estimates are based on the predicted numbers anaemic in each risk group given by the mean and individual-level standard deviations at each realisation of the Bayesian multivariate model for Hb, and are shown for two demographic groups by Province.

Population Attributable Risk fractions by Province, % (95% BCI)						
	Central	Coast	Eastern	Nyanza	Rift Valley	Western
14 year old girls:						
Observed prevalence of anaemia	3.3	37.3	34.7	30.8	19.4	25.0
PAF for:						
Med/high intensity hookworm	<i>n/a</i>	13.0 (3.7, 23.1)	5.9 (0, 12.5)	8.2 (0, 30.0)	5.0 (0, 12.5)	29.0 (1.4, 55.1)
Malaria	<i>n/a</i>	5.8 (2.0, 10.4)	<i>n/a</i>	27.6 (0, 49.1)	<i>n/a</i>	23.6 (0, 40.9)
Schistosomiasis	<i>n/a</i>	3.1 (0, 0.08)	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
Acute malnutrition ^a	<i>n/a</i>	18.5 (0, 37.9)	18.4 (0, 73.0)	<i>n/a</i>	16.0 (0, 52.5)	<i>n/a</i>
7 year old boys:						
Observed prevalence of anaemia	5.0	43.3	23.2	43.2	22.4	19.2
PAF for:						
Med/high intensity hookworm	<i>n/a</i>	5.6 (1.0, 10.3)	<i>n/a</i>	11.0 (0, 21.6)	3.4 (0, 6.6)	13.5 (0, 27.5)
Malaria	<i>n/a</i>	1.5 (0.3, 6.1)	<i>n/a</i>	14.3 (0, 32.0)	<i>n/a</i>	15.8 (0, 28.1)
Schistosomiasis	<i>n/a</i>	0 (0, 5.5)	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
Acute malnutrition ^a	<i>n/a</i>	16.4 (0, 33.5)	16.4 (0, 67.1)	<i>n/a</i>	14.1 (0, 45.9)	<i>n/a</i>

^a Exposure to “acute malnutrition”, prevalence of wasting in under 5s exceeds 5%
n/a – no children are exposed (i.e. prevalence of predictor is zero in this demographic group)

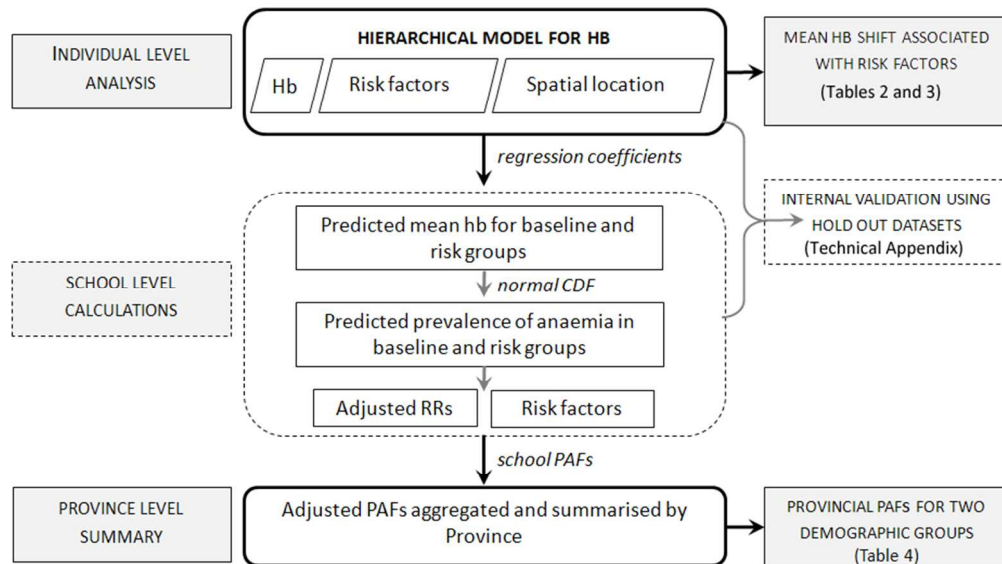


Figure 1. Schematic of adopted analysis strategy. The dual aim of this analysis strategy was to (i) model the impact of risk factors on children's haemoglobin (Hb) levels, and (ii) estimate the adjusted Population Attributable Risk Fraction (PAF) associated with important, modifiable risk factors. This process was repeated at each realisation of the Bayesian hierarchical model to provide posterior mean estimates and credible intervals for all parameters of relevance.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

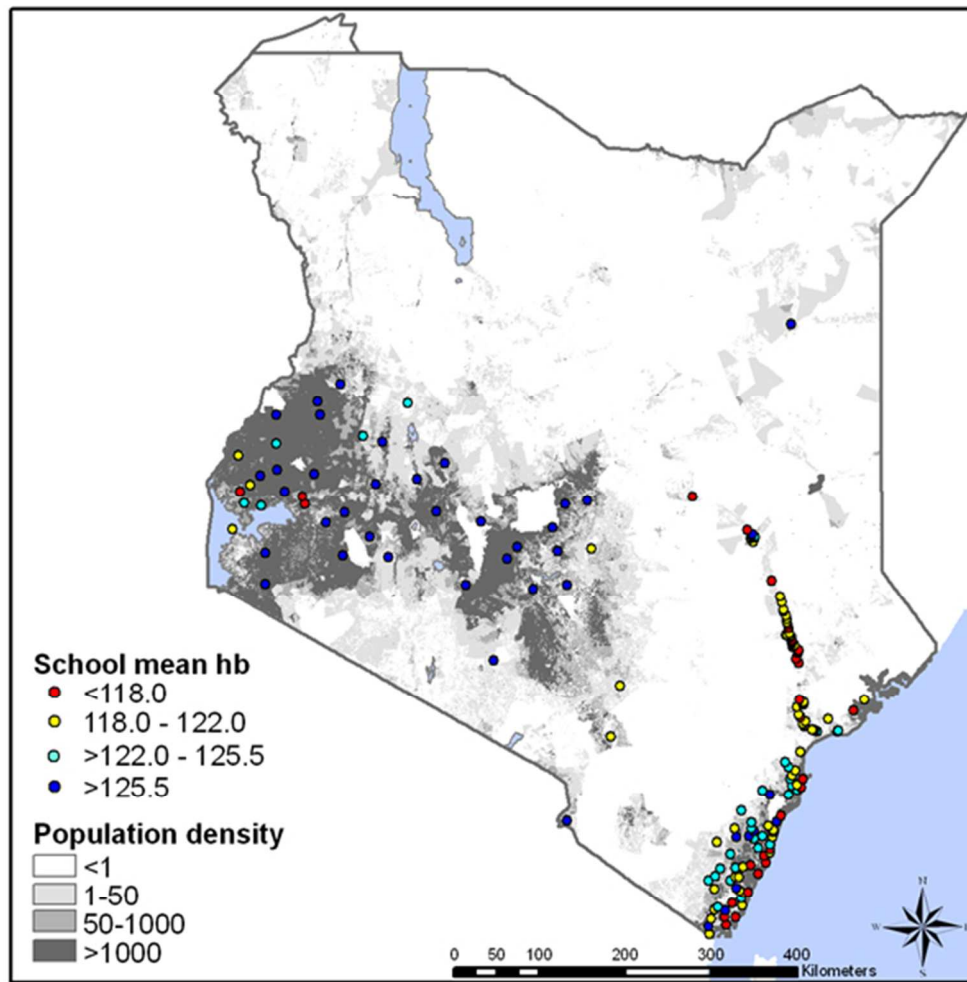


Figure 2 Spatial distribution of mean haemoglobin distribution across Kenya, by school. HB is adjusted for altitude, population density (persons / km²) is shown for reference.

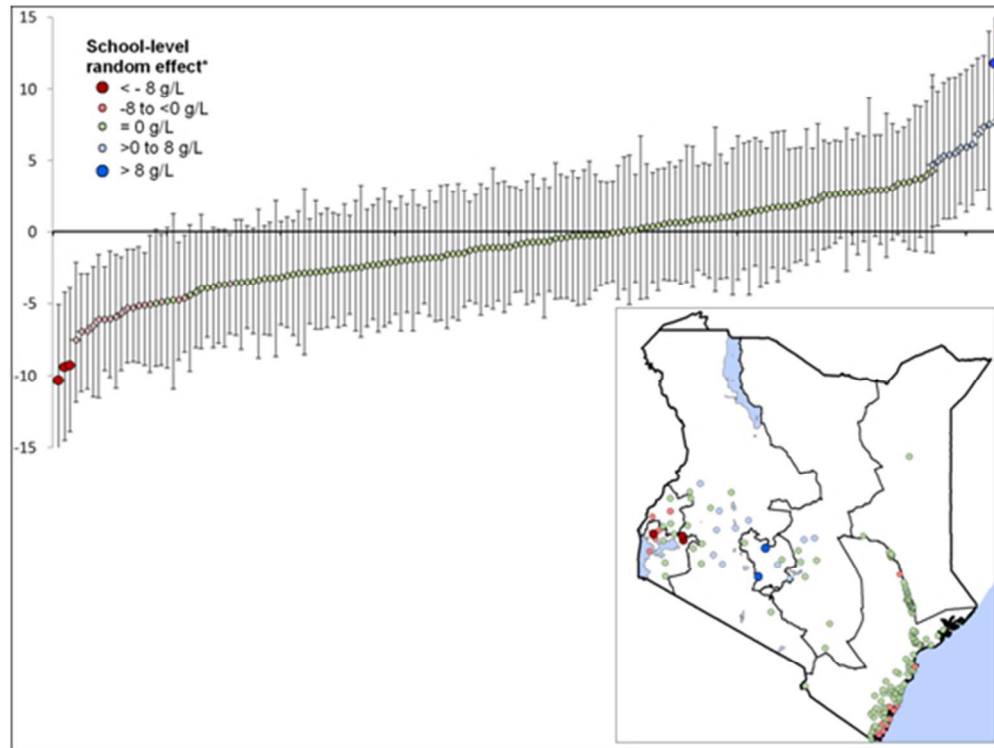


Figure 3. Graph of spatial random-effect (school-level residual) from the multivariate, Bayesian hierarchical model for Hb. Each point represents one school; the posterior mean is shown by the dot and the 95% BCI are represented by the error bars. The horizontal straight line in the middle of the graph represents the overall mean of the residuals, which are zero centred. Inset shows the school locations, shaded by the value of the school-level residual.

SUPPLEMENTARY TECHNICAL APPENDIX

Estimating the relative contribution of parasitic infections and nutrition for anaemia among school-aged children in Kenya: a subnational geostatistical analysis

Model fit for all Bayesian estimation was carried out in WinBUGs 1.4.1 (Imperial College London, and MRC, London, UK). Semi-variogram analysis was conducted in R, and all other descriptive statistical analysis and data management were carried out in Stata 11.0 software (StataCorp)

Bayesian Kriging of Nutritional Proxy Data

Proxy data for macro-nutritional status in school-aged children were derived from the Kenya 2008-09 Demographic and Health Survey (DHS). This dataset includes anthropometric measurements for 5,470 children aged under 6 years and 7,548 women of child-bearing age (ages 15 to 49). These were chosen to provide a general indication of the degree of malnutrition prevalent in the communities where school-children were surveyed. As DHS clusters rarely coincide with the location of surveyed schools (see Figure S1), it was first necessary to interpolate cluster-level prevalence to school locations using a weighted moving average approach (kriging). In order to provide a realistic account of uncertainty in both the predicted prevalence and the covariance functions used to define weighting, thus providing a range of potential prevalence values, this was embedded in a Bayesian framework using a simple binomial model structure.

The numbers of examined (n_i) and positive (Y_i) individuals for each anthropometric indicator in cluster i ($i=1, \dots, N$) were modelled as binomial variates in the form:

$$Y_i \sim \text{Binomial}(n_i, p_i)$$

$$\text{logit}(p_i) = \alpha + u_i$$

1 where α is the global mean prevalence and u_i a geostatistical random effect modelled using an isotropic,
 2 stationary exponential decay function: $f(d_{ab}; \phi) = \exp[-(\phi d_{ab})]$ where d_{ab} is the straight-line
 3 distance between cluster locations a and b, and ϕ is the rate of decline of spatial correlation. A non-
 4 distance between cluster locations a and b, and ϕ is the rate of decline of spatial correlation. A non-
 5 informative normal prior was used for α , the prior distribution of ϕ was uniform with upper and lower
 6 bounds set at 0.05 and 50 and the precision of u_i was given non-informative gamma distribution.
 7
 8 Following a burn-in of 10,000 iterations, the model was run for a further 10,000 iterations with thinning
 9 every ten iterations, during which predictions were stored.
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

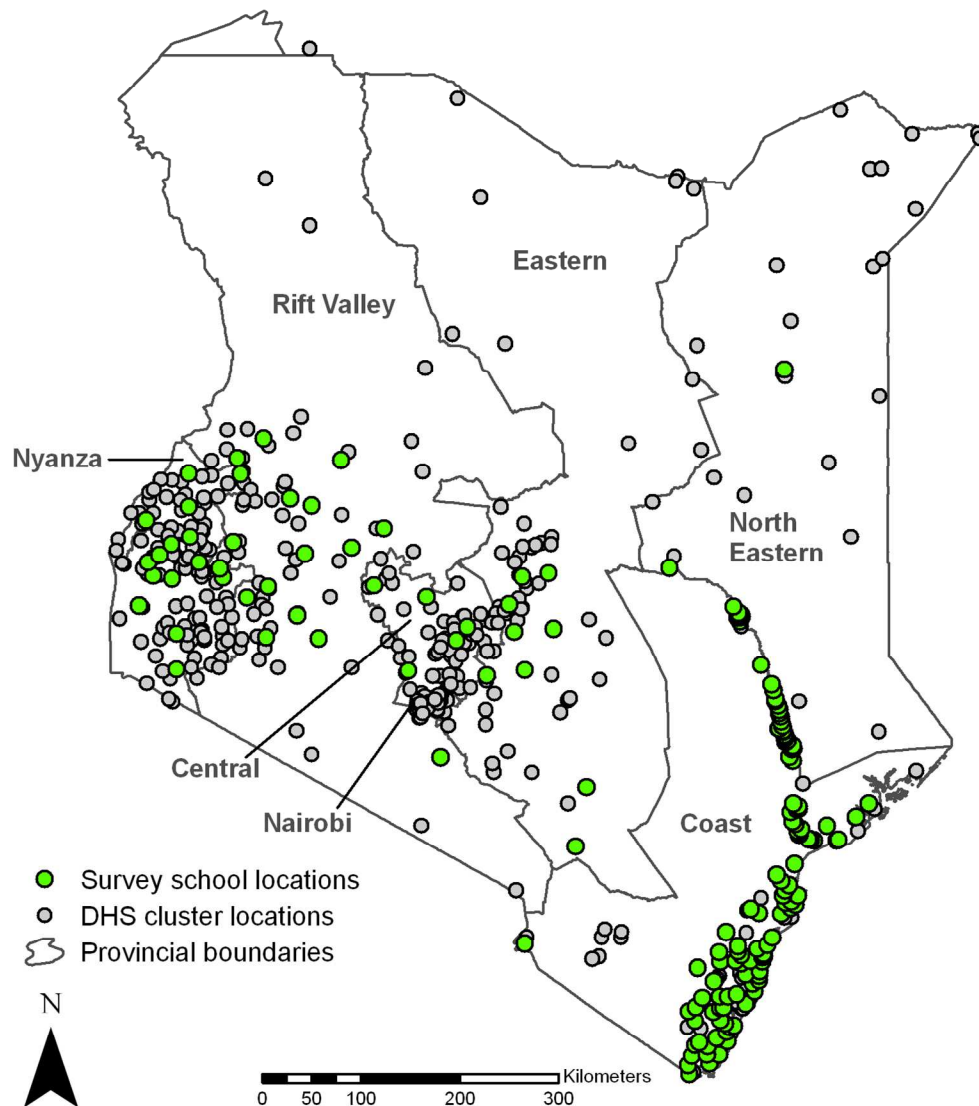


Figure S1: Location of 167 school surveys and 400 DHS cluster sites across Kenya. School survey data was collected between 2008-2010, and DHS data between 2008-2009. Provincial boundaries are shown for reference.

Table S1: Summary of results from Bayesian kriging of anthropometric indicator data collected during 2008-2009 Kenya DHS survey.

Anthropometric Indicator	Province	Mean observed prevalence at DHS clusters, % (range)	Mean predicted prevalence at school survey sites, % (range)
Wasting (in children < 5 yrs)	Central	3.3 (0, 25.0)	4.1 (3.9, 4.7)
	Coast	8.5 (0, 35.3)	10.5 (4.8, 22.8)
	Eastern	7.2 (0, 35.3)	5.1 (4.1, 6.4)
	North Eastern	12.6 (0, 33.3)	19.0 (12.7, 22.1)
	Nyanza	2.1 (0, 11.8)	2.5 (2.0, 3.7)
	Rift Valley	4.9 (0, 21.1)	4.1 (2.5, 9.1)
	Western	2.0 (0, 23.1)	2.3 (1.9, 3.6)
Low BMI (in women 15-49 yrs)	Central	8.5 (0, 4.0)	9.8 (5.6, 19.0)
	Coast	11.7 (0, 44.4)	16.8 (7.2, 28.5)
	Eastern	16.8 (0, 47.6)	13.6 (5.5, 23.9)
	North Eastern	27.6 (6.2, 52.9)	21.0 (16.3, 31.4)
	Nyanza	7.4 (0, 31.3)	7.2 (4.6, 9.7)
	Rift Valley	14.8 (0, 55.0)	11.8 (5.6, 22.3)
	Western	7.4 (0, 21.7)	8.2 (7.1, 9.3)
Consumed iron rich foods in past 24 hours (children < 5 yrs)	Central	14.8 (0, 45.5)	17.7 (9.1, 25.0)
	Coast	16.2 (0, 55.0)	17.5 (6.8, 31.4)
	Eastern	16.3 (0, 50.0)	18.1 (16.0, 20.7)
	North Eastern	15.3 (0, 60.7)	9.3 (5.9, 15.6)
	Nyanza	19.2 (0, 55.6)	21.4 (13.9, 27.7)
	Rift Valley	17.3 (0, 55.6)	17.2 (7.5, 23.8)
	Western	31.7 (0, 69.2)	28.3 (24.2, 35.3)

As can be seen from Table S1, observed prevalence of each indicator in the DHS dataset varied considerably within each Province. This is reflected in part by the range of predicted values shown for the school survey sites, although as geostatistical models tend to regress towards the mean the range in predicted values is considerably less. As under- and over-prediction are unlikely to be biased, this suggests that true association between proxies of malnutrition and hb levels in school-aged children are likely to be underestimated.

Bayesian hierarchical model of haemoglobin

For this model, haemoglobin level for each child i ($i = 1, \dots, N$) in school j ($j = 1, \dots, N$) is considered a normal variable outcome Y_{ij} in the form:

$$Y_{i,j} \sim \text{Normal}(\mu_{i,j}, \sigma^2)$$

$$\mu_{i,j} = \alpha + \sum_{k=1}^K \beta_k X_{i,j,k} + \sum_{l=1}^L \beta_l X_j + \beta_f u_j + \beta_g v_j + w_j$$

where α is the intercept, $\sum_{k=1}^K \beta_k X_{i,t,k}$ the matrix of individual-level covariates, $\sum_{l=1}^L \beta_l X_j$ the matrix of school-level covariates, u_j is considered a dummy variable labelled $u_j=1$ when prevalence of wasting is predicted to be medium/high severity (5-15%) and v_j a dummy variable labelled $v_j=1$ when prevalence of wasting is predicted to be extremely high (>15%) and w_j is a geostatistical random effect modelled using an isotropic, stationary exponential decay function: $f(d_{ab}; \phi) = \exp[-(\phi d_{ab})]$ where d_{ab} is the straight-line distance between pairs of school locations a and b , and ϕ is the rate of decline of spatial correlation. Non-informative priors were used for α and the coefficients (normal prior with mean 0 and precision 1 x 106), the prior distribution of ϕ was uniform with upper and lower bounds set at 1 and 50 and the precision of v_j was given a non-informative gamma distribution. Values of u_j and v_j at each model iteration are defined by a random variable Z_j , which is given an informative beta prior defined using the posterior mean and standard deviation of predicted prevalence of wasting (described above):

$$Z_j \sim \text{Beta}(\alpha_j, \beta_j)$$

$$0.05 > Z_j > 0.15, \quad u_j=1$$

$$0.15 > Z_j > 0.15, \quad v_j=1$$

Following a burn-in of 9,000 iterations, the values for the intercept and coefficients were stored for 1,000 iterations and model convergence was assessed using diagnostic tests for convergence and by visually inspecting the time series plots. Convergence was successfully achieved after 10,000, and the model was

run for a further 10,000 iterations with thinning every ten iterations, during which prediction locations were stored.

Model output therefore consisted of samples from the posterior distribution hb levels for individuals and summarised by school (mean Hb).

Model adjusted Population Attributable Risk Fractions (PAFs)

In general, analysis of adjusted PAFs for binary outcomes such as anaemia rely on using the adjusted odds ratios (ORs) resulting from binomial GLMs. Whilst regional PAFs can be estimated using this approach based on the spatial distribution of each risk factor by region, it cannot take into account spatial variation in the underlying haemoglobin levels of resident children and thus assumes the impact of each risk factor is constant. For this reason, we chose to use a two step process that explicitly considers variation in underlying haemoglobin levels to estimate PAFs for anaemia by Province across Kenya, as outlined below.

At each realisation of the Bayesian hierarchical model for haemoglobin, the mean hb level (μ_j) was estimated at each school for a baseline group (7 year old boys and 14 year old girls, in socio-economic quantiles 2-4) and for each risk group included in Table 4 (main text), including school-level covariates (SF status) and the geostatistical random error term (w_j). The normal cumulative distribution function (CDF) was then used to estimate the prevalence of anaemia (i.e. the proportion of children with hb levels lower than age- and sex-standardised cut-offs) for baseline and risk groups based on predicted μ_j and the modelled within-school standard deviation σ^2 , and the adjusted relative risk (RR) for each risk factor estimated. In order to produce internally valid estimates when using adjusted RRs, the method proposed by Kelinbaum *et al*⁽⁵²⁾ (see⁽³¹⁾ for discussion) was used to estimate the adjusted PAFs:

$$PAF = \sum_{i=0}^k pd_i \left(\frac{RR_i - 1}{RR_i} \right)$$

where pd_i is the proportion of cases falling into the i th exposure level, and is calculated as the observed school-level prevalence in either boys aged 6-8 years or girls aged 13-15, and RR_i is the relative risk the i th

1 exposure level with the unexposed group. School-level PAFs were then aggregated and summarised by
 2 Province. Model output therefore consisted of samples from the posterior distribution of Province mean
 3 PAFs for each risk factor.
 4
 5
 6
 7
 8
 9

10 **Model Validation**

11
 12 To assess the predictive performance of the final model of Hb, and associated predicted prevalence of Hb,
 13 hold-out validation data subsets were generated by random division of the data into quintiles. The
 14 Bayesian geostatistical model was then trained on four subsets whilst simultaneously predicting
 15 haemoglobin for the fifth, excluded subset. This was repeated for each subset, giving and observed and
 16 predicted Hb value for all surveyed individuals. The ability of the model to predict known mean Hb (and
 17 associated prevalence of anaemia) was assessed by three validation statistics: mean prediction error
 18 (which provides a measure of the bias of the predictor); absolute mean prediction error (which provides a
 19 measure of the accuracy of the predictor); and the correlation coefficient (providing a measure of
 20 association between observed and predicted Hb values).
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34

35 Results of this model validation process are shown below in Table S2. When summarised by school, there
 36 was a very high degree of correlation between observed and expected mean Hb. Mean error and absolute
 37 mean error of the anaemia predictions was -6%, suggesting that the model consistently over-predicted
 38 prevalence. At an individual level the model was less successful, on average producing a 10 g/L
 39 discrepancy between predicted and observed Hb levels. This difference between school and individual
 40 level predictive validity is reflective of the large residual variation between individuals within schools
 41 indicated by the model (Table 2).
 42
 43
 44
 45
 46
 47
 48
 49
 50

51 **Table S2:** Measures of correlation, bias and accuracy of Bayesian hierarchical regression models at individual and school
 52 levels.

53 Outcome	54 Mean error	55 Absolute mean error	56 Correlation coefficient
57 Individual-level:			
58 Hb level, g/L	- 0.005	+ 10.15	0.392
59 Anaemia	+ 0.042	+ 0.364	0.298
60 School-level:			
Mean Hb, g/L	- 1.74	+1.79	0.982

1 Prevalence of anaemia, % - 0.061 +0.065 0.951

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For peer review only



Estimating the relative contribution of parasitic infections and nutrition for anaemia among school-aged children in Kenya: a subnational geostatistical analysis

Journal:	<i>BMJ Open</i>
Manuscript ID:	bmjopen-2012-001936.R1
Article Type:	Research
Date Submitted by the Author:	22-Jan-2013
Complete List of Authors:	Pullan, Rachel; London School of Hygiene & Tropical Medicine, Department of Disease Control Gitonga, Carol; Kenya Medical Research Institute-Wellcome Trust Research Programme, Mwandawiro, Charles; Eastern and Southern Africa Centre of International Parasite Control, Kenya Medical Research Institute (KEMRI), Snow, Robert; Kenyan Medical Research Institute/University of Oxford-Wellcome Trust Research Programme, Malaria Public Health & Epidemiology Group Brooker, Simon; London School of Hygiene and Tropical Medicine,
Primary Subject Heading:	Epidemiology
Secondary Subject Heading:	Global health, Haematology (incl blood transfusion), Nutrition and metabolism
Keywords:	NUTRITION & DIETETICS, Epidemiology < INFECTIOUS DISEASES, Public health < INFECTIOUS DISEASES, Tropical medicine < INFECTIOUS DISEASES

1
2 **Estimating the relative contribution of parasitic infections and nutrition for anaemia**
3
4 **among school-aged children in Kenya: a subnational geostatistical analysis**
5
6

7 Rachel L Pullan¹, Carol Gitonga^{1,2}, Charles Mwandawiro³, Robert W Snow^{2,4} and Simon J Brooker^{1,2}
8
9

10 ¹ *London School of Hygiene and Tropical Medicine, London, UK*

11 ² *Kenya Medical Research Institute-Wellcome Trust Research Programme, Nairobi, Kenya;*

12 ³ *Eastern and Southern Africa Centre of International Parasite Control, Kenya Medical Research Institute*
13 *(KEMRI), Nairobi, Kenya*

14 ⁴ *Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK*
15
16
17

18 **Correspondence:** rachel.pullan@lshtm.ac.uk; Faculty of Infectious and Tropical Diseases, London School
19
20
21 of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objectives: To quantify geographical variation in the relative contribution of parasitic infections, socio-economic factors and malnutrition in the aetiology of anaemia among school children across Kenya, thereby providing a rational basis for the targeting of an integrated school health package.

Design: Nationally representative cross-sectional survey data were collected using standard protocols. For all included children, data were recorded on haemoglobin concentration and common parasitic infections (*Plasmodium falciparum*, hookworm and schistosomes) and socio-economic indicators. Ecological proxies of malnutrition and food security were generated using Demographic and Health Survey and UN Food and Agriculture Organization food security data respectively. Spatially explicit, multilevel models were used to quantify impact upon child haemoglobin concentration.

Setting: Randomly selected schools in ecologically diverse settings across Kenya.

Main outcome measures: Mean haemoglobin concentration adjusted for infection, nutritional and socio-economic risk factors; associated risk ratios and adjusted Population Attributable Fractions (PAFs) for anaemia, by region.

Results: Data were available for 16,941 children in 167 schools; mean Hb was 122.1 g/L and 35.3% of children were anaemic. In multivariate analysis, mean Hb was significantly lower in boys and younger children. Severe malnutrition and interactions between *P. falciparum* and hookworm infections were significantly associated with lower Hb, with greater impacts seen for co-infected children. The contribution of risk factors to anaemia risk varied by province: in 14 year old girls PAFs ranged between 0-27.6% for *P. falciparum*, 0-29.0% for hookworm and 0-18.4% for severe malnutrition.

Conclusions: The observed geographical heterogeneity in the burden of anaemia attributable to different aetiological factors has important implications for the rational targeting of anti-anaemia interventions that can be included into an integrated school health programme.

Article Summary

Article Focus

- Cross-sectional haematological and parasitological school survey data for the period 2008-2010 in Kenya
- Spatially-explicit, Bayesian multi-level models quantifying impact of infection, nutrition and social risk factors on child haemoglobin concentration and adjusted Population Attributable Fractions for anaemia by province.
- Implications of spatial heterogeneity in underlying haemoglobin status and risk factors for cost-effective geographical targeting of school-based anti-anaemia interventions, including deworming, malaria control and school feeding.

Key messages

For school-aged children living in Kenya:

- *P. falciparum*, hookworm and schistosomes are all significantly associated with lower mean Hb, with the largest impact seen for coinfection with hookworm and *P. falciparum*, emphasising the importance of parasitic infections in the aetiology of anaemia in African school-age populations
- Poor food intake is suggested to be an important contributor to anaemia in school-aged children.
- There is pronounced geographical heterogeneity in the burden of anaemia attributable to these identified risk factors underscoring importance of appropriately targeting interventions

Strengths and limitations of the study

- Large and uniquely detailed dataset coupled with novel Bayesian modelling framework that allowed for simultaneous quantification of the mean shift in Hb associated with each risk factor, and sub-national risk of anaemia in school-aged children attributable to risk factors.
- Dependence upon cross-sectional design and ecological data limits inference of causality; incorporation of individual, nutritional-related factors and iron status indicators would improve quantification of their involvement.

Background

Children of school-going age (5-14 years) make up one quarter of the population of sub-Saharan Africa,¹ and the proportion of these children attending school has never been higher, in large part due to recent global and national commitments to universal primary education school enrolment.² Yet many children enrolled in school suffer from poor health and malnutrition and as a result are less likely to attend and complete school or achieve their full educational potential.³ For example, up to 40% of 5-14 year olds are estimated to be anaemic,^{4,5} which can affect cognition and learning.^{6,7} Anaemia also affects the health of older girls as they reach puberty, potentially impacting upon future motherhood and the nutrition of the next generation of learners.^{8,9} Fortunately, a number of simple and cost-effective interventions are available to avert the burden of anaemia, including malaria control,^{10,11} deworming,¹² iron supplementation^{13,14} and school-feeding¹⁵ – interventions which are increasingly delivered as part of an integrated school health package.¹⁶

Identifying the optimal combination of school health interventions requires reliable information on both the geographical distribution of anaemia and the relative contribution of different aetiological factors. A previous study of anaemia among preschool children in West Africa employed a geostatistical framework to interpolate estimates of anaemia among children aged 1-4 years collected through demographic health surveys (DHS), whilst incorporating ecological predictions of malaria and helminth infection. Whilst this work helped define the geographical distribution and limits of anaemia in the region it did not quantify spatial variation in the risk attributable to different aetiological factors.¹⁷ Here we take advantage of uniquely detailed haematological data collected during a comprehensive, national school health survey conducted across Kenya between 2008 and 2010.¹⁸ An initial analysis showed that the relative contribution of *Plasmodium falciparum* to anaemia varied according to malaria transmission zones.¹⁹ We extend this to investigate the relationships between haemoglobin concentration (Hb) and common parasitic infections (including *P. falciparum*, hookworm and schistosomes), socio-economic indicators and proxies of malnutrition and food security for school children across Kenya. We use an innovative Bayesian approach that quantifies the mean shift in Hb associated with each risk factor, which we use to estimate

1 sub-national variation in the risk of anaemia in school-aged children attributable to malnutrition, malaria
2
3 and helminths to define targets for integrated packages of school-health interventions.
4
5
6
7

8 **Methods**

9 ***School survey data:***

10
11
12
13
14 This study uses cross-sectional school survey data collected between September 2008 and March 2010
15
16 using standard protocols, described elsewhere.^{18, 20} In brief, 480 government primary schools were
17
18 selected to provide a national assessment of malaria and anaemia among school children. In a random
19
20 subset of 167 schools, stool and urine samples were also collected to assess prevalence and intensity of
21
22 helminth infections. At each school, 10 children plus 1 reserve of each sex were randomly selected from
23
24 each of classes 2-6 from the children present that day. This captures children aged 4 to 16, although 80%
25
26 of included children were aged between 8 and 13 years. Selected children were asked to provide a
27
28 fingerprick blood sample, which was used to assess haemoglobin (Hb) using a portable photometer
29
30 (Haemocue, Angelholm, Sweden) and *Plasmodium* infection in the peripheral blood based on microscopy-
31
32 corrected malaria rapid diagnostic test (RDT)¹⁸ results, whereby slides were subsequently read for all RDT
33
34 positives. Children in all schools were asked to provide stool samples which were examined for the eggs of
35
36 intestinal nematodes (*Ascaris lumbricoides*, *Trichuris trichiura* and hookworm species) and *S. mansoni*
37
38 using the Kato-Katz method. In schools on the coast, where *Schistosoma haematobium* is widespread,
39
40 children were asked to provide urine samples which were examined by urine filtration. A questionnaire
41
42 was administered to all pupils to obtain data on insecticide-treated net ownership and use, history of
43
44 anthelmintic medication, key socioeconomic variables, household construction and education of the
45
46 child's parent/guardian. School locations were determined using a hand-held Global Positioning System
47
48 and classified as urban using an updated 2010 urban extents (UE) mask derived from GRUMP.^{21, 22}
49
50
51
52
53
54

55 ***Nutritional proxy data***

1 Anthropometric and food intake data were not available for all of our study population, and we instead
2
3 relied on ecological measures. Proxy data for macro-nutritional status in school-aged children were
4
5 derived from the Kenya 2008-09 Demographic and Health Survey (DHS).^{23, 24} Anthropometric
6
7 measurements were extracted for children ≤ 5 years and for women of child-bearing age (ages 15 to 49)
8
9 (see Table 1). For children ≤ 5 years, the reported proportion consuming food rich in iron (flesh and organ
10
11 meat) in the past 24 hours was also extracted. These indicators, which provide a general measure of
12
13 nutritional status in the wider community, were geo-located to the centre of each sampling cluster.
14
15 Bayesian kriging²⁵, a geostatistical interpolation method that accounts for the error introduced by
16
17 estimating the semivariogram model (a function of the variability in outcome against the distance
18
19 separating observation points), was used to interpolate a distribution of possible values for each indicator
20
21 to survey school locations. This was done by incorporating the full posterior distribution of the Bayesian
22
23 semivariogram model, estimated using the Bayesian statistical software WinBUGS version 14.1 (Medical
24
25 Research Council Biostatistics Unit and Imperial College London). Mean predicted prevalence values were
26
27 then categorised according to WHO-defined cut-off values for severity of malnutrition by prevalence
28
29 range (Table 1). See Technical Appendix for further details.
30
31
32
33
34
35

36 Two additional indicators were used to describe general food availability at each school location:
37
38 predominant source of livelihood (based on 2010 FAO Livelihood zones^{26, 27}); and food security conditions
39
40 at the time of the survey (based on the FAO Food Insecurity Severity Scale^{27, 28}). Both indicators were
41
42 obtained from the Famine Early Warning System Network (FEWS NET; <http://www.fews.net>). Information
43
44 on the intended provision of formal, government-run school-feeding programmes was abstracted from
45
46 the Homegrown School Meals Programme (HGSMP) and the Emergency HGSMP 2010 registers provided
47
48 by the Kenya Ministry of Education (*personal communication*).
49
50
51
52
53
54

55 **Analysis**

56
57
58
59
60

1 A schematic summarising the major steps in our analysis strategy is provided in Figure 1. Haemoglobin
2 levels were adjusted for altitude and individuals were classified as anaemic using age and sex specific cut-
3 offs.²⁹ Children were classified as having medium/high intensity hookworm infection if infection intensity
4 exceeded 100 epg, and medium/high intensity *S. haematobium* infection if infection intensity exceeded
5 500 epg (both representing the top 10th percentile). Reported information on ownership of household
6 assets and household construction was used to generate wealth indices using principle component
7 analysis,³⁰ and resulting scores were divided into quintiles.

8
9
10
11
12
13
14
15
16
17
18
19 The initial candidate set of predictor variables for univariate analysis included age, sex, relative socio-
20 economic status, education-level of the household head, protective behaviours (bednet usage and
21 deworming) and infection status with hookworm, schistosome and *P. falciparum*, in addition to school-
22 level ecological covariates including urban/rural location, predicted prevalence of anthropometric
23 indicators in younger children and women of reproductive age, livelihood zone and food security status
24 and intended provision of school-feeding. In order to select candidate variables for spatial multivariable
25 analysis, univariate analysis was undertaken using logistic regression (criteria: Wald test $P < .1$) using Stata
26 11.1 software (StataCorp). Standard errors were adjusted for dependence between individuals within
27 schools using a clustered sandwich estimator.³¹ Backward-stepwise elimination was then used to generate
28 a minimum adequate multivariate model; excluded covariates ($P > .05$) were retested in the minimal
29 model to confirm lack of association. Basic age-sex interaction terms were also investigated, but there
30 was insufficient evidence that these improved the overall fit of the model.

31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47 Retained covariates were included in a spatial multivariable Bayesian mixed-effects linear regression
48 model using WinBUGs 14.1. This model included individual and school-level ecological fixed effects and a
49 school-level geostatistical random effect (using an isotropic, stationary exponential decay function),²⁵
50 thereby incorporating a degree of spatial smoothing. Uncertainty in the predicted prevalence of
51 anthropometric proxy indicators was incorporated by modelling these covariates as random variables

1 with a beta distribution, defined using the prediction posterior distribution. At each model realisation,
2
3 malnutrition was categorised as either low, medium/high or severe for each school location, according to
4
5 Table 1.
6
7
8
9

10 Given the large observed differences in mean Hb and prevalence of predictors by region and demographic
11
12 group, adjusted Population Attributable Fractions (PAFs) for anaemia were estimated for two indicator
13
14 demographic groups (boys aged 7 years, and girls aged 14 years) for each province in Kenya. In brief, at
15
16 each model realisation we estimated the prevalence of anaemia in each of the risk groups based upon the
17
18 posterior mean and standard deviation of the Hb distribution within each school. These were used to
19
20 estimate (i) the relative risk of anaemia for each risk group compared with baseline risk, and (ii) the
21
22 associated adjusted PAF, as described by Rockhill *et al* (1998).³² Estimates were subsequently
23
24 summarised by province. Full statistical notation of Bayesian geostatistical modelling, PAF and model
25
26 validation procedures are presented in Technical Appendix. Model validation using hold out datasets,
27
28 presented in detail in the accompanying Technical Appendix, suggest very high degree of correlation at
29
30 the school level between observed and expected mean Hb levels ($R^2 > 0.98$), and prevalence of anaemia
31
32 ($R^2 > 0.95$) suggesting the approach we used to estimate anaemia prevalence (and associated attributable
33
34 fractions) was appropriate.
35
36
37
38
39
40
41
42
43

44 **Results:**

45
46 Data on Hb, malaria parasitaemia and helminth infection were available for 16,941 children (aged 5-16) in
47
48 167 schools (Figure 2). Mean altitude-adjusted Hb concentration was 122.1 g/L, with 35.3% of children
49
50 estimated to be anaemic. The spatial distribution of the unadjusted Hb data is presented in Figure 2,
51
52 summarised by school. There were significant geographical heterogeneity in mean Hb across the country
53
54 (Moran's I $p < 0.001$), with the lowest average levels observed in Coast Province (120.1 g/L) and the highest
55
56 in Central Province (134.0 g/L). Univariate analysis suggested that Hb levels increased with age and with
57
58
59
60

1 relative socio-economic status, and were significantly lower in individuals infected with *P. falciparum* or
2
3 medium/high intensity helminth infections (Table 1). Hb was also significantly lower in areas with
4
5 predicted severe malnutrition, although there was no ecological relationship with predicted consumption
6
7 of iron rich foods. Univariate analysis also suggests a relationship between Hb and food availability, with
8
9 significantly lower Hb levels in children living in very food insecure areas.
10

11 ***Bayesian hierarchical model of haemoglobin***

12
13
14
15
16
17 Table 3 presents a multivariate analysis of factors associated with Hb levels. On average, boys had lower
18
19 mean Hb compared with girls by 1.53 g/L (95% BCI: 1.13-1.94 g/L), and older children tended to have
20
21 higher Hb levels, increasing 1.08 g/L (95% BCI: 0.99-1.17 g/L) each yearly age interval. There was strong
22
23 evidence for a multiplicative negative interaction between infection status and Hb levels, highlighting the
24
25 health impacts of parasitic infections in school-aged children. Children with medium/high intensity
26
27 hookworm infection had Hb levels 1.03 g/L (95% BCI: 0.02-2.00 g/L) lower than uninfected children, those
28
29 with *P. falciparum* 3.00 g/L (95% BCI: 1.68-4.29 g/L) lower and *S. haematobium* infection 1.75 g/L (95%
30
31 BCI: 0.60-2.92 g/L); children co-infected with both *P. falciparum* and medium/high intensity hookworm
32
33 infection had Hb levels 5.56 g/L (95% BCI: 2.67-8.64 g/L) lower than uninfected children.
34
35
36
37
38
39

40 Proxy measures of nutritional status were also strongly associated with Hb levels. School-children living in
41
42 areas predicted to have severe prevalence of wasting had Hb levels that were on average 9.74 g/L (95%
43
44 BCI: 7.45-12.16 g/L) lower than those living in low risk areas. School-feeding programmes also appear to
45
46 have been targeted to districts with increased risk of anaemia, although examination of the 95% credible
47
48 intervals suggests no statistical difference within target districts between mean Hb levels in schools with
49
50 and without intended provision of formal school-feeding programmes. Perhaps surprisingly, food security
51
52 at the time of the survey (a qualitative measure that allows adjustment for acute conditions and the
53
54 impact of seasonality) was not associated with Hb levels in multivariate analysis. Similarly, after
55
56
57
58
59
60

1 accounting for all other factors in the model, relative socio-economic status was not significantly
2 associated with Hb levels.
3
4
5
6
7

8
9 The random effects variance components from this model indicate that most of the residual variation
10 remains between children within schools: of the total unexplained variability in Hb, 13.8% occurred at the
11 school level, whilst 86.2% occurred between children within schools. Closer examination of the spatial
12 school-level random effect suggests that, after accounting for included predictors, most schools (81%) had
13 similar Hb levels that could not be distinguished statistically (shown in green in Figure 3). There was
14 however some residual spatial clustering of schools with higher/lower than expected Hb levels (foci radii
15 of ~190 km, Table 2) with significantly lower mean Hb levels observed in schools in the far south of Coast
16 Province and in Nyanza and Western provinces (Figure 3). Interestingly, comparison of observed
17 prevalence of *P. falciparum* and medium/high intensity hookworm suggests that prevalence of both
18 infections was higher for these schools than for other schools in these provinces (*P. falciparum* Kruskal-
19 Wallis p 0.02; medium/high intensity hookworm Kruskal-Wallis p 0.01). In contrast, mean Hb levels were
20 significantly higher for schools in Rift Valley, Central and Eastern provinces
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 **Adjusted Population Attributable Risk Fractions**

38
39 Using the final Bayesian hierarchical model, we estimated for two demographic groups (girls aged 14
40 years, and boys aged 7 years) the PAF of anaemia due to *P. falciparum*, medium/high intensity hookworm
41 infection, medium/high intensity *S. haematobium* infection, and acute malnutrition (prevalence of
42 wasting in under 5s exceeding 15%) for each province in Kenya, adjusting for all other included predictors
43 (Table 4). Our results suggest that the relative contribution of each risk factor varies considerably by
44 province. For example, in Coast Province 13% (95% BCI: 4-23%) of anaemia in 14 year old girls attending
45 school may be attributable to medium/high intensity hookworm, comparing with 6% (95% BCI: 2-10%) to
46 *P. falciparum* infection and 3.1% (95% BCI: 0-8%) to schistosomiasis. Meanwhile, results suggest that
47 reduction in acute malnutrition in Coast Province (to <5% wasting in ≤5s) would reduce anaemia
48
49
50
51
52
53
54
55
56
57
58
59
60

1 prevalence by 19% (95% BCI: 0-38%). This contrasts with Western Province, where the estimated risk of
2 anaemia attributable to medium and high intensity hookworm is 29% (95% BCI: 1-55%) and to *P.*
3 *falciparum* 24% (95% BCI 0-41%); no school children in this district are exposed to *S. haematobium* and
4 acute malnutrition in under 5s is predicted not to exceed 5%. The wide 95% BCIs observed reflect (i) the
5 smaller sample sizes for provinces other than Coast (given that N = schools, not individuals), and (ii) the
6 additional uncertainty surrounding proxies for malnutrition. Trends in PAFs observed for predictors in 7
7 year old boys mirror those for older girls but are generally lower, a reflection of the lower baseline Hb
8 levels in this group.
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23

24 Discussion

25
26 As global commitment and resources for school health and nutrition initiatives grow,¹⁶ there is an urgent
27 need to strengthen our understanding of the likely impact of interventions aimed at tackling important
28 health and development outcomes such as anaemia. This requires a detailed understanding of the relative
29 contribution of aetiological factors at policy relevant scales. Here we present a uniquely detailed national
30 analysis of risk factors for anaemia among schoolchildren in 167 schools across Kenya. We introduce a
31 new approach using multilevel linear modelling to provide a detailed description the varying contribution
32 of risk factors by administrative unit (province), whilst fully accounting for spatial variation in underlying
33 mean Hb. Results provide a vital decision-making tool for cost-effective targeting the delivery of
34 integrated, packages of interventions tailored to each province, with the aim of reducing anaemia in
35 schoolchildren. For example, school-feeding programmes in Coast, Eastern and Rift Valley provinces may
36 want to consider including iron-rich food stuffs, whilst programmes in Nyanza and Western provinces
37 omitting malaria control initiatives may not see desired improvements in childhood anaemia.
38
39
40
41
42
43
44
45
46
47
48
49
50

51
52
53
54
55 Our analysis found evidence that *P. falciparum*, hookworm and schistosomes were each significantly
56 associated with lower mean Hb, observations consistent with current knowledge. Such findings
57
58
59
60

1 corroborate recent results from the Global Burden of Disease study, which suggests that globally for
2 anaemia 10.2% of years lived with a disability (YLD) across all ages could be attributed to hookworm,
3 17.9% to malaria and 3.4% to schistosomiasis.³³ Our analysis provides three additional insights into the
4 epidemiology of anaemia. First, we confirm small-scale studies³⁴⁻³⁸ that have suggested that coinfection
5 with hookworm and *P. falciparum* may have a multiplicative effect on Hb level – effects which occur
6 across Kenya. Second, we demonstrate pronounced geographical heterogeneity in the burden of anaemia
7 attributable to identified risk factors. Associations for hookworm and *P. falciparum* were most
8 pronounced in Western Province, where results suggest that together these two infections were
9 attributable for close to 53% (95% BCI: 1.4-96.0%) of anaemia cases in girls aged 14 years. Together, these
10 findings emphasise the importance of parasitic infections in the aetiology of anaemia in school-aged
11 children and underscore current efforts to control helminths and malaria as part of integrated school
12 health programmes.^{39,40} The PAFs provided here are useful for translating our findings into numbers that
13 can help policy decision makers appreciate the potential benefits of carefully targeted programmes
14 tailored to each province that strike the best balance between dietary interventions, deworming and
15 malaria control.^{17,41}

16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36 Third, whilst acknowledging that we rely on proxy indicators, our results do suggest that poor food intake
37 is an important contributor to anaemia in school children: the final model indicates that in communities
38 where acute malnutrition is a severe public health problem (i.e. where wasting in under 5s exceeds 15%),
39 mean Hb levels in school-aged children are on average 9.7 g/L lower than well nourished communities,
40 and as such eliminating acute malnutrition as a public health problem could avert close to 20% of anaemia
41 cases in 14 year old girls. This is consistent with limited evidence from smaller individual-level studies
42 suggesting an association between under-nutrition and anaemia in school-aged populations.^{42,43}

43
44
45
46
47
48
49
50
51
52
53
54 Our spatially explicit hierarchical approach seeks to maximise the usefulness of this observational, cross-
55 sectional data and results show that whilst much of the between school variation could be explained by
56
57
58
59
60

1 included risk factors, a high degree of variation does still remain within schools. This suggests that there
2 are individual-level factors beyond those measured that still need to be considered. In addition to
3 individual iron status or food intake there are a number of candidates. HIV-infection is potential likely to
4 be a contributor, as even in asymptomatic HIV infected individuals anaemia is a common manifestation.⁴⁴
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

⁴⁵ We were also unable to account for genetic traits such as sickle-cell and other haemoglobinopathies,⁴⁶
or for the potential effect of menarche in adolescent girls.⁴⁷

A number of additional study limitations should also be acknowledged. As we have previously
acknowledged,²⁰ cross-sectional studies such as this are inevitably subject to a number of limitations
including the potential of ecological fallacy, especially when incorporating population-level estimates. We
acknowledge that our indicator of school-feeding (inclusion in the government school-feeding registers) is
far from optimal, and the study would have benefitted from detailed individual-level data on iron status
and food intake. In addition, modelled associations are likely to be lower than the true impact, due to
non-systematic measurement error in our infection indicators, which rely on single samples using
standard parasitological diagnosis with <100% sensitivity.^{48, 49} Despite a robust stratified random sampling
frame, our results may also be subject to some selection bias, as very ill children such as those with acute
infections and severe chronic malnutrition may have been absent on the day of the school survey and
thus not included in the study sample. Although contemporary regional or school-level attendance data is
not available, overall school attendance rates are known to vary geographically across Kenya.⁵⁰ For
example, in 2005 (the most recent year for which sub-national net primary school attendance figures are
available) reported net attendance in Nyando District (Nyanza Province) was greater than 90% for both
males and females, whilst in Turkana District (Rift Valley Province) levels were between 30-40%.⁵⁰ If Hb
levels differ systematically between non-attending and attending school-aged children, this may act to
either dilute or exaggerate geographical variation in Hb levels

1 An additional factor that would influence individual-level variation is the onset of menarche in girls: no
2 specific data were collected to record this, and inclusion of a basic age-sex interaction term did not
3 improve model fit suggesting that more detailed information is required. Finally, this analysis is unable to
4 account for the impact of recently cleared infections on current Hb status. and the presented results
5 represent a snapshot of the period 2008-2010 and it is conceivable that seasonality and intervention-
6 related factors may have confounded results. However, the inclusion of mosquito net use, recent
7 deworming or food security (a seasonal covariate) in multivariate models did not improve model fits,
8 suggesting that temporal and control factors were relatively unimportant.
9
10
11
12
13
14
15
16
17
18
19
20

21 **Conclusions**

22 A recent focus on school-child and adolescent health has emphasised the need to develop a deeper
23 understanding of factors impacting health and development of this population group.⁵¹⁻⁵⁴ Here we
24 present a novel two-step process to model Hb levels and then predict associated anaemia prevalence in
25 different areas of Kenya. Our approach adds considerable value over and above what could be achieved
26 modelling anaemia directly, and the adopted geostatistical Bayesian framework allows for propagation of
27 uncertainty associated at each step (including generation of predicted ecological covariates) into the final
28 confidence limits. Future expansion of this model could look at interpolating existing results to make
29 predictions in non-surveyed areas, although this would require generation of values for all included
30 individual-level covariates for each prediction location, increasing prediction uncertainty. Our results
31 highlight the marked geographical heterogeneity in the burden of anaemia attributable to identified risk
32 factors in Kenya, which has major implications for effective targeting of packages of school-based
33 interventions. We suggest that deworming can help tackle anaemia both in western and coastal Kenya,
34 whereas malaria control is likely to have the greatest impact in western Kenya, whilst school-feeding
35 programmes can help reduce anaemia in Coast, Eastern and Rift Valley provinces. The priority now should
36 be to explore in detail the health and educational impact of carefully designed integrated infection and
37 nutritional interventions in this age group.
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

We are very grateful to all the teachers, parents and especially children who participated in this survey and to the many facilitators and members of the survey teams who collected the data and to Don Bundy for comments. The survey work was supported by the Wellcome Trust UK as part of fellowship support to SJB (#81673) and by the Division of Malaria Control, Ministry of Public Health and Sanitation through a grant from the Department for International Development (DfID) through the WHO Kenya Country Office to Dr AM Noor and Dr E Juma.

Footnotes

Contributors RLP was responsible for designing and conducting analysis, interpretation, drafting and production of the final manuscript. CG contributed to the survey design, data assembly and cleaning and contributed to final manuscript. CM and RWS contributed to survey design, interpretation of results and contributed to final manuscript. SJB designed the field surveys, provided scientific guidance and contributed to interpretation and preparation of the final manuscript. All authors read and approved the final manuscript.

Funding The fieldwork was supported by the Wellcome Trust UK and by the Division of Malaria Control, Ministry of Public Health and Sanitation. RLP acknowledges support of the Bill & Melinda Gates Foundation and CWG is supported by a Commonwealth Scholarship from the DfID. SJB is supported by a Wellcome Trust Senior Fellowship in Basic Biomedical Science (098045) and RWS is supported by a Wellcome Trust Principal Fellowship (079080).

Competing interests None.

Data sharing statement: The school survey dataset used for this analysis is available for download through the Global Atlas of Helminth Infection (www.thiswormyworld.org). The WinBUGs statistical code is available from the corresponding author at rachel.pullan@lshtm.ac.uk.

References

1. United Nations Department of Economic and Social Affairs. World population prospects, 2010 revision; Available from: <http://esa.Un.Org/unpd/wpp/>.
2. Global Partnership for Education. All children learning: Report 2011. Global Partnership for Education; Available from: <http://allchildrenlearningreport.globalpartnership.org>
3. World Bank. Fresh: Focusing resources on education and school health. A fresh start to improving the quality and equity of education. UNESCO, UNICEF, WHO & World Bank; 2000
4. McLean E, Cogswell M, Egli I, et al. Worldwide prevalence of anaemia, WHO vitamin and mineral nutritional information system, 1993-2005. *Public Health Nutrition*. 2008;12(4):444-454.
5. Best C, Neufingerl N, van Geel L, et al. The nutritional status of school-aged children: Why should we care? *Food and Nutrition Bulletin*. 2010;31(3):400-417.
6. Grantham-McGregor S, Ani C. A review of studies on the effect of iron deficiency on cognitive development in children. *Journal of Nutrition*. 2001 Feb;131(2S-2):649S-666S.
7. Fretham SJ, Carlson ES, Georgieff MK. The role of iron in learning and memory. *Advances in Nutrition*. 2011;2(2):112-121.
8. Scholl TO. Iron status during pregnancy: Setting the stage for mother and infant. *American Journal of Clinical Nutrition*. 2005;81(5):1218S-1222S.
9. Dharmalingam A, Navaneetham K, Krishnakumar CS. Nutritional status of mothers and low birth weight in india. *Maternal and Child Health Journal*. 2010;14(2):290-298.
10. Clarke SE, Jukes M, Njagi JK, et al. Effect of intermittent preventive treatment of malaria on health and education in schoolchildren: A cluster-randomised, double-blind, placebo-controlled trial. *Lancet*. 2008;372(9633):127-138.
11. Leenstra T, Phillips-Howard PA, Kariuki SK, et al. Permethrin-treated bed nets in the prevention of malaria and anemia in adolescent schoolgirls in western kenya. *American Journal of Tropical Medicine and Hygiene*. 2003;68(S86-S93).
12. Smith JL, Brooker S. The impact of hookworm infection and deworming on anaemia in non-pregnant populations: A quantitative update. *Tropical Medicine and International Health*. 2010;15(7):776-795.
13. Leenstra T, Kariuki SK, Kurtis JD, et al. The effect of weekly iron and vitamin a supplementation on hemoglobin levels and iron status in adolescent schoolgirls in western kenya. *European Journal of Clinical Nutrition*. 2009;63(2):173-182.
14. Gera T, Sachdev HP, Nestel P. Effects of iron supplementation on physical performance in children and adolescents: Systematic review of randomized controlled trials. *Indian Pediatrics*. 2007;44:15-24.
15. Bundy DAP, Burbano C, Grosh M, et al. *Rethinking school feeding: Social safety nets, child development and the education sector*. The World Bank: New York; 2009.
16. Bundy DAP. *Rethinking school health: A key component of education for all*. Washington: The World Bank; 2012.
17. Soares Magalhães RJ, Clements ACA. Mapping the risk of anaemia in preschool-age children: The contribution of malnutrition, malaria, and helminth infections in West Africa. *PLoS Medicine*. 2011;8(6):e1000438.
18. Gitonga C, Karanja P, Kihara J, et al. Implementing school malaria surveys in kenya: Towards a national surveillance system. *Malaria Journal*. 2010;9(1):306.
19. Gitonga CW, Edwards T, Karanja PN, et al. *Plasmodium* infection, anaemia and mosquito net use among school children across different settings in Kenya. *Tropical Medicine and International Health*. 2012;17(7):858-870.
20. Brooker S, Pullan RL, Gitonga CW, et al. *Plasmodium*-helminth coinfection and its sources of heterogeneity across east africa. *Journal of Infectious Diseases* 2012;205(5):841-852.
21. Balk DL, Deichmann U, Yetman G, et al. Determining global population distribution: Methods, applications and data. *Advances in Parasitology*. 2006;62:119-156.

22. Balk D, Pozzi F, Yetman G, et al. The distribution of people and the dimension of place: Methodologies to improve the global estimation of urban extents. Urban Remote Sensing Conference; 2004; Available at: <http://www.iussp.org/Activities/wgc-urb/balk.pdf>
23. Kenya National Bureau of Statistics, Measure DHS. Kenya: Dsh, 2008-09 - final report. Calverton, Maryland, USA: Measure DHS, ICF Macro; 2010
24. Measure DHS. Kenya 2008-09 DHS data. Available from: <http://www.measuredhs.com>
25. Diggle P, Moyeed R, Tawn J. Model-based geostatistics (with discussion). *Applied Statistics*. 1998;47:299-350.
26. FEWSNET. Famine early warning system network (FEWSNET) livelihood zone map for Kenya. FEWSNET and USAID; 2011
27. Grillo J, Holt J. Application of the livelihood zone maps and profiles for food security analysis and early warning: USAID FEWSNET; 2009
28. FEWSNET. Famine early warning system network (FEWSNET) Kenya archive: Food security update. FEWSNET and USAID; 2011
29. WHO. Iron deficiency anaemia: Assessment, prevention and control. A guide for programme managers. Geneva: World Health Organisation; 2001
30. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data--or tears: An application to educational enrollments in states of india. *Demography*. 2001;38(1):115-32.
31. Williams RL. A Note on Robust Variance Estimation for Cluster-Correlated Data. *Biometrics*. 2000; 56: 645-6.
32. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *American Journal of Public Health*. 1998;88(1):15-19.
33. Jasrasaria J, Murray CJL, Naghavi M. Global Burden of Disease 2010 Estimate: Strategy Report for Anaemia Envelope. [Unpublished report]. *Personal Communication* 2012.
34. Brooker S, Akhwale WS, Pullan R, et al. Epidemiology of *Plasmodium*-helminth coinfection in Africa: Potential impact on anaemia and prospects for combining control. *American Journal of Tropical Medicine & Hygiene*. 2007;77:88-98.
35. Pullan R, Brooker S. The health impact of polyparasitism in humans: Are we under-estimating the burden of parasitic diseases? *Parasitology*. 2008:1-12.
36. Nkuo-Akenji TK, Chi PC, Cho JF, et al. Malaria and helminth co-infection in children living in a malaria endemic setting of Mount Cameroon and predictors of anaemia. *Journal of Parasitology*. 2006;92(6):1191-1195.
37. Midzi N, Mtapuri-Zinyowera S, Mapingure MP, et al. Consequences of polyparasitism on anaemia among primary school children in Zimbabwe. *Acta Trop*. 2010;115:103-111.
38. Ezeamama EE, McGarvey ST, Acosta LP, et al. The synergistic effect of concomitant schistosomiasis, hookworm, and trichuris infections on children's anemia burden. *PLoS Negl Trop Dis*. 2008;2(6):e245.
39. Bundy DAP, Shaeffer S, Jukes M, et al. School based health and nutrition programs. *Chapter 58 in "Disease Control Priorities for Developing Countries"*. 2006.
40. Brooker S, Clarke S, Snow RW, et al. Malaria in african schoolchildren: Options for control. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2008;102(4):304-305.
41. Rowe AK, Powell KE, Flanders WD. Why population attributable fractions can sum to more than one. *American Journal of Preventative Medicine*. 2004;26:243-249.
42. Anumudu C, Adolami M, Igwe C, et al. Nutritional anaemia and malaria in pre-school and school-aged children. 2008;1(11-17).
43. Daboné C, Delise HF, Receveur O. Poor nutritional status of schoolchildren in urban and peri-urban areas of Ouagadougou (Burkina Faso). *Nutrition Journal*. 2011;10:34.
44. Calis JC, van Hensbroek MB, de Haan RJ, et al. HIV-associated anemia in children: A systematic review from a global perspective. *AIDS*. 2008;22(10):1099-1112.
45. Levine AM, Berhane K, Masri-Lavine L, et al. Prevalence and correlates of anemia in a large cohort of HIV-infected women: Women's interagency HIV study. *Journal of Acquired Immune Deficiency Syndrome*. 2001;26:28-35.
46. Tolentino K, Friedman JF. An update on anaemia in less developed countries. *American Journal of Tropical Medicine and Hygiene*. 2007;77:44-51.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
47. Leenstra T, Kakiuki SK, Kurtis JD, et al. Prevalence and severity of anaemia and iron deficiency: Cross-sectional studies in adolescent schoolgirls in western Kenya. *European Journal of Clinical Nutrition*. 2004;58(4):681-691.
48. Booth M, Vounatsou P, N'Goran E K, et al. The influence of sampling effort and the performance of the kato-katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Cote d'Ivoire. *Parasitology*. 2003;127(Pt 6):525-31.
49. Okell LC, Ghani AC, Lyons E, et al. Submicroscopic infection in *Plasmodium falciparum* endemic populations: A systematic review and meta-analysis. *Journal of Infectious Disease*. 2009;15(10):1509-1517.
50. World Bank. World development report 2007: Development and the next generation. Washington: The World Bank
51. Education and Policy Data Center. Kenya 2005 District Primary Education Profiles. Available From: www.epdc.org/country/kenya
52. Little AW, Green A. Successful globalisation, education and sustainable development. *International Journal of Education and Development*. 2009;29:166-174.
53. Sawyer SM, Afifi RA, Bearinger LH, et al. Adolescence: A foundation for future health. *The Lancet*. 2012;379(9826):1630-1640.
54. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiological research. Belmont, CA: Lifetime Learning Publications; 1982. p. 163.

Figure legends:

Figure 1 Schematic of adopted analysis strategy. The dual aim of this analysis strategy was to (i) model the impact of risk factors on children's haemoglobin (Hb) levels, and (ii) estimate the adjusted Population Attributable Risk Fraction (PAF) associated with important, modifiable risk factors. This process was repeated at each realisation of the Bayesian hierarchical model to provide posterior mean estimates and credible intervals for all parameters of relevance.

Figure 2 Spatial distribution of mean haemoglobin distribution across Kenya, by school. HB is adjusted for altitude, population density (persons / km²) is shown for reference.

Figure 3 Graph of spatial random-effect (school-level residual) from the multivariate, Bayesian hierarchical model for Hb. Each point represents one school; the posterior mean is shown by the dot and the 95% BCI are represented by the error bars. The horizontal straight line in the middle of the graph represents the overall mean of the residuals, which are zero centred. Inset shows the school locations, shaded by the value of the school-level residual.

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Tables:

Table 1 Definition of included anthropometric indicators, and population-based cut-off values for severity of malnutrition, as defined by the World Health Organization.

Anthropometric Measurement		Severity of malnutrition (prevalence range for < 6 years)		
		Low severity	Medium/high severity	Very high severity
Stunting	Height-for-age Z-score > -2 SD	<20%	20-40%	>40%
Wasting	weight-for-height Z-score > -2 SD	<5%	5-15%	>15%
Underweight	weight-for-age <i>or</i> BMI-for-age Z-score > -2 SD	<10%	10-30%	>30%

Table 2 Mean Hb concentration according to infection status, nutritional proxies and other characteristics in 16,941 children aged 5-16 years (univariate analysis).

Variable	Data type and source	Children No (%)	Mean Hb shift, g/L (SE)	p
Sex				
Female	Individual – KHS	8,314 (49.1)	(vs. female)	
Male	Individual – KHS	8,627 (50.9)	- 1.36 (0.28)	<0.001
Age (in years)	Individual – KHS	-	+ 0.89 (0.11)	<0.001
Socio-economic indicators				
Educated household head ¹	Individual – KHS	8,333 (49.2)	+ 0.82 (0.41)	0.05
Household owns a mobile phone	Individual – KHS	10,065 (59.4)	+ 1.42 (0.40)	<0.001
SES quintile				
Lowest	Individual – KHS	3,272 (20.4)	(vs. lowest)	
Mid (quintiles 2-4)		9,926 (61.8)	+ 2.35 (0.50)	<0.001
Highest		2,854 (17.8)	+ 3.07 (0.82)	<0.001
Urban location	School-level – digital UE map	1,237 (7.3)	1.56 (1.52)	0.3
Current infections				
<i>P. falciparum</i> parasitaemic	Individual – KHS	607 (3.6)	-3.36 (0.77)	<0.001
Med/high intensity hookworm (>100 epg)	Individual – KHS	931 (5.5)	-2.81 (0.69)	<0.001
Med/high intensity <i>S. haematobium</i> (>500 epg)	Individual – KHS	799 (4.7)	-3.81 (1.02)	<0.001
Usually sleeps under a net	Individual – KHS	9,763 (57.6)	- 1.78 (0.52)	0.001
Taken anthelmintic in previous 6 months	Individual – KHS	7,494 (44.2)	+ 1.24 (0.48)	0.01
Nutritional proxies				
Wasting in under 5s (%) ²				
Under 5%	Ecological – interpolated from 2008 DHS	4,027 (23.8)	(vs <5%)	
5-15%		10,257 (60.6)	-5.58 (1.04)	<0.001
Over 15%		2,657 (15.7)	-8.14 (1.21)	<0.001
Low BMI in adult women (%) ³				
Under 20%	Ecological – interpolated from 2008 DHS	12,867 (76.0)	(vs. <20%)	
Over 20%		4,072 (24.0)	-1.98 (0.80)	0.01
Consumption iron rich foods in under 5s (%) ⁴				
Under 25%	Ecological – interpolated from 2008 DHS	14,572 (86.0)	(vs. <25%)	
Over 25%		2,369 (14.0)	+ 1.07 (1.24)	0.4
School-feeding (SF) status: ⁵				
Not a SF district	School-level – MoE registers	3,690 (21.8)	(vs. not SF dist)	
SF district – no programme in school		8,417 (49.7)	-7.34 (1.08)	<0.001
SF district – programme in school		4,834 (28.5)	-4.80 (1.19)	<0.001
Generalised livelihood zone ⁶				
Mixed farming		3,295 (19.5)	(vs. mixed)	
Pastoral and agro-pastoral	Ecological – FEWSNET	4,991 (29.5)	+ 1.11 (1.45)	0.4
Riverine and fishing		641 (3.8)	+ 2.08 (2.42)	0.4
Medium potential farming		4,073 (24.0)	- 2.93 (1.13)	0.02
High potential farming		3,941 (23.3)	+ 1.37 (2.14)	0.2
Food security status ⁷				
Generally food secure	Ecological – FEWSNET	2,756 (16.3)	(vs. secure)	
Moderately food insecure		7,755 (45.8)	-3.20 (1.27)	0.01
Highly food insecure		6,048 (35.7)	- 6.17 (1.25)	<0.001
Extremely food insecure		382 (2.3)	- 5.12 (1.52)	0.001

KSH; Kenya School Health survey (conducted 2008-2010). DHS; Kenya Demographic Health Survey conducted 2008/2009. MoE; Kenya Ministry of Education. FEWSNET; Famine early warning system network. SE; standard error.

¹ household head educated to primary complete or above (reported).

^{2,3,4} Interpolated from Kenya 2008/2009 DHS survey

² school included in either the Homegrown School Meals Programme (HGSM) or Emergency HGSM 2010 registers

⁶ based on the USAID/FEWSNET Generalised Livelihood Zone 2010 Classification for Kenya

⁷ based on the USAID/FEWSNET Food Insecurity Severity Scale for the quarter corresponding to the KSH survey date.

Table 3 Bayesian hierarchical multivariate regression model for haemoglobin levels.

Variable	Mean Hb shift (g/L)	95% BCI / SD
Baseline haemoglobin	117.3	(113.9, 120.5)
Sex (male)	-1.53	(-1.94, -1.13)
Age (in years)	+1.08	(+0.99, +1.17)
SES quintile (vs. quintiles 2-4)		
lowest	-0.56	(-1.14, +0.02)
highest	0.25	(-0.31, +0.86)
Med/high intensity <i>S. haematobium</i> (>500 epg)	-1.75	(-2.92, -0.60)
Med/high intensity hookworm (>100 epg)	-1.03	(-2.00, -0.02)
<i>P. falciparum</i> parasitaemic	-3.00	(-4.29, -1.68)
Med/high hookworm – <i>P. falciparum</i> co-infection	-5.56	(-8.64, -2.67)
Wasting in under 5s (vs <5% prevalence)		
5-15%	-4.39	(-6.56, -2.09)
>15%	-9.74	(-12.16, -7.45)
School-feeding (SF) status:		
SF district – no programme in school	-2.14	(-4.76, +0.46)
SF district – programme in school	-2.78	(-5.43, -0.21)
Random effect terms		
Individual σ^2	183.6	(2.02)
Spatial σ^2 (school-level)	29.5	(7.93)
Range of spatial correlation (km)	194	(99,318)

Table 4: Predicted adjusted population attributable risk fractions for key predictors across Kenya. Estimates are based on the predicted numbers anaemic in each risk group given by the mean and individual-level standard deviations at each realisation of the Bayesian multivariate model for Hb, and are shown for two demographic groups by Province.

Population Attributable Risk fractions by Province, % (95% BCI)						
	Central	Coast	Eastern	Nyanza	Rift Valley	Western
14 year old girls:						
Observed prevalence of anaemia	3.3	37.3	34.7	30.8	19.4	25.0
PAF for:						
Med/high intensity hookworm	<i>n/a</i>	13.0 (3.7, 23.1)	5.9 (0, 12.5)	8.2 (0, 30.0)	5.0 (0, 12.5)	29.0 (1.4, 55.1)
Malaria	<i>n/a</i>	5.8 (2.0, 10.4)	<i>n/a</i>	27.6 (0, 49.1)	<i>n/a</i>	23.6 (0, 40.9)
Schistosomiasis	<i>n/a</i>	3.1 (0, 0.08)	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
Acute malnutrition ^a	<i>n/a</i>	18.5 (0, 37.9)	18.4 (0, 73.0)	<i>n/a</i>	16.0 (0, 52.5)	<i>n/a</i>
7 year old boys:						
Observed prevalence of anaemia	5.0	43.3	23.2	43.2	22.4	19.2
PAF for:						
Med/high intensity hookworm	<i>n/a</i>	5.6 (1.0, 10.3)	<i>n/a</i>	11.0 (0, 21.6)	3.4 (0, 6.6)	13.5 (0, 27.5)
Malaria	<i>n/a</i>	1.5 (0.3, 6.1)	<i>n/a</i>	14.3 (0, 32.0)	<i>n/a</i>	15.8 (0, 28.1)
Schistosomiasis	<i>n/a</i>	0 (0, 5.5)	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
Acute malnutrition ^a	<i>n/a</i>	16.4 (0, 33.5)	16.4 (0, 67.1)	<i>n/a</i>	14.1 (0, 45.9)	<i>n/a</i>

^a Exposure to “acute malnutrition”, prevalence of wasting in under 5s exceeds 5%
n/a – no children are exposed (i.e. prevalence of predictor is zero in this demographic group)

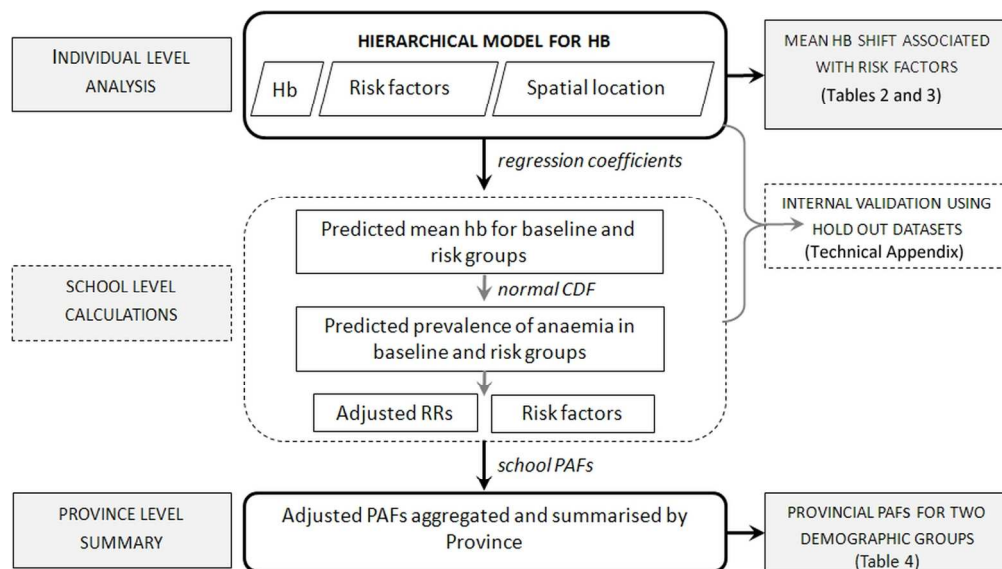


Figure 1. Schematic of adopted analysis strategy. The dual aim of this analysis strategy was to (i) model the impact of risk factors on children’s haemoglobin (Hb) levels, and (ii) estimate the adjusted Population Attributable Risk Fraction (PAF) associated with important, modifiable risk factors. This process was repeated at each realisation of the Bayesian hierarchical model to provide posterior mean estimates and credible intervals for all parameters of relevance.

160x90mm (300 x 300 DPI)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

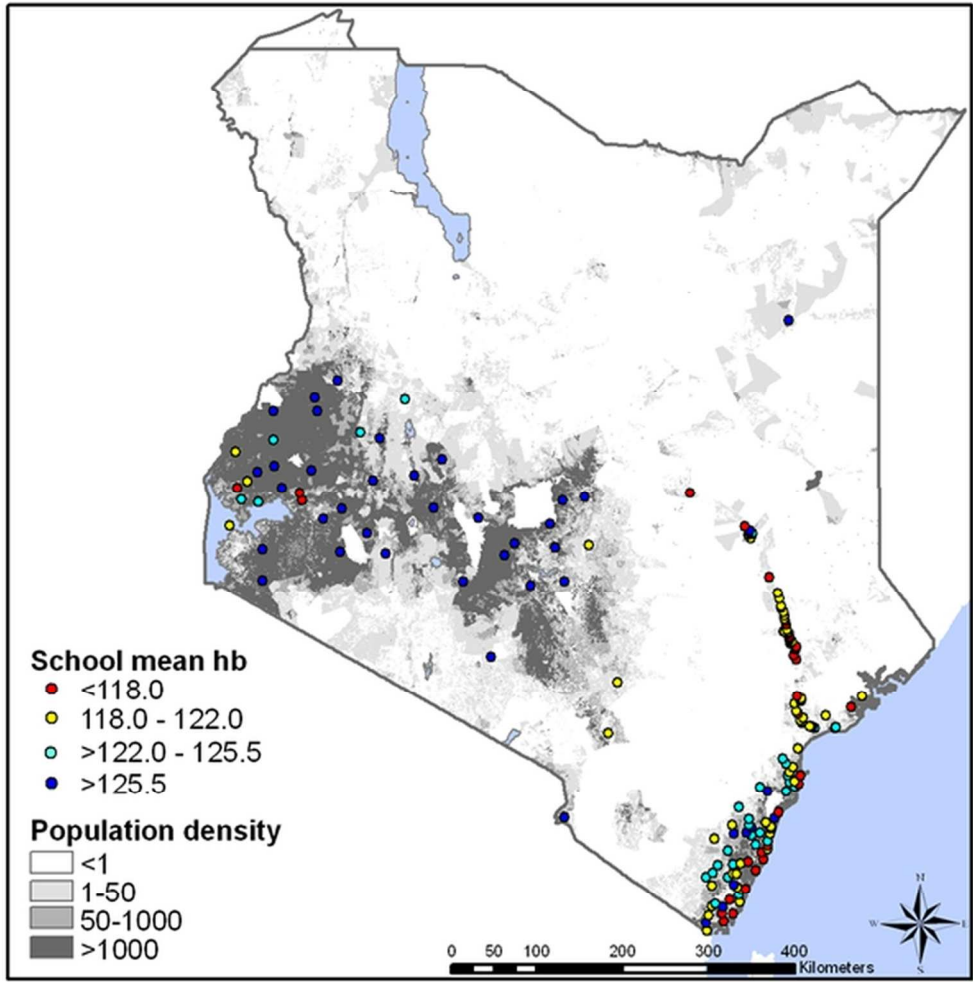


Figure 2 Spatial distribution of mean haemoglobin distribution across Kenya, by school. HB is adjusted for altitude, population density (persons / km²) is shown for reference.
90x90mm (300 x 300 DPI)

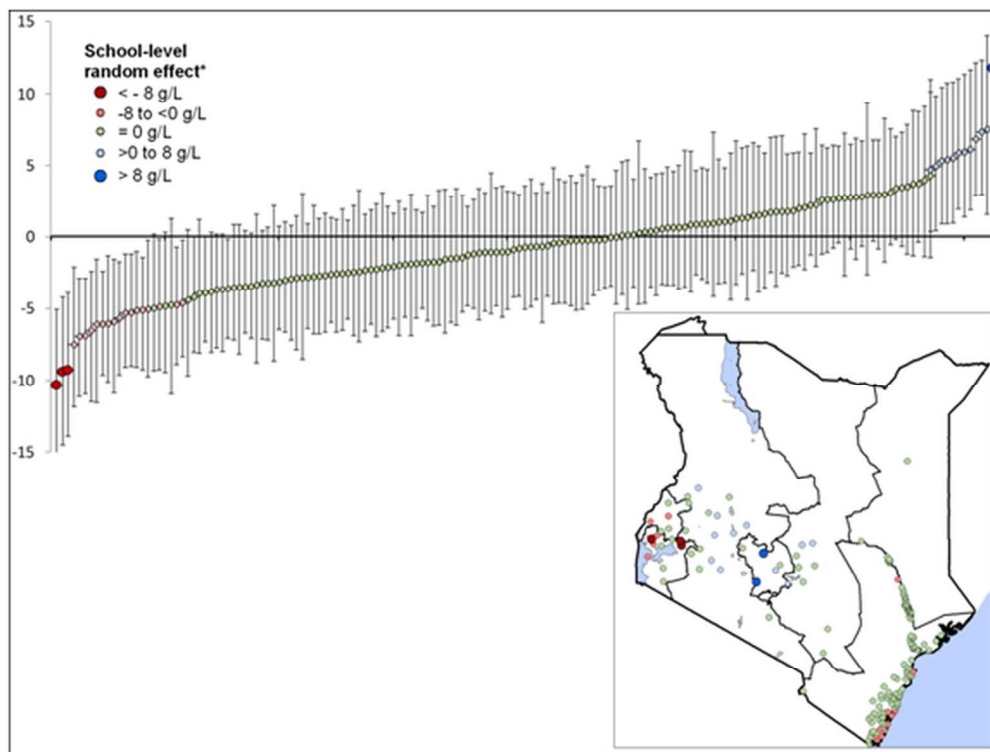


Figure 3. Graph of spatial random-effect (school-level residual) from the multivariate, Bayesian hierarchical model for Hb. Each point represents one school; the posterior mean is shown by the dot and the 95% BCI are represented by the error bars. The horizontal straight line in the middle of the graph represents the overall mean of the residuals, which are zero centred. Inset shows the school locations, shaded by the value of the school-level residual.
119x90mm (300 x 300 DPI)

SUPPLEMENTARY TECHNICAL APPENDIX**Estimating the relative contribution of parasitic infections and nutrition for anaemia among school-aged children in Kenya: a subnational geostatistical analysis**

Model fit for all Bayesian estimation was carried out in WinBUGs 1.4.1 (Imperial College London, and MRC, London, UK). Semi-variogram analysis was conducted in R, and all other descriptive statistical analysis and data management were carried out in Stata 11.0 software (StataCorp)

Bayesian Kriging of Nutritional Proxy Data

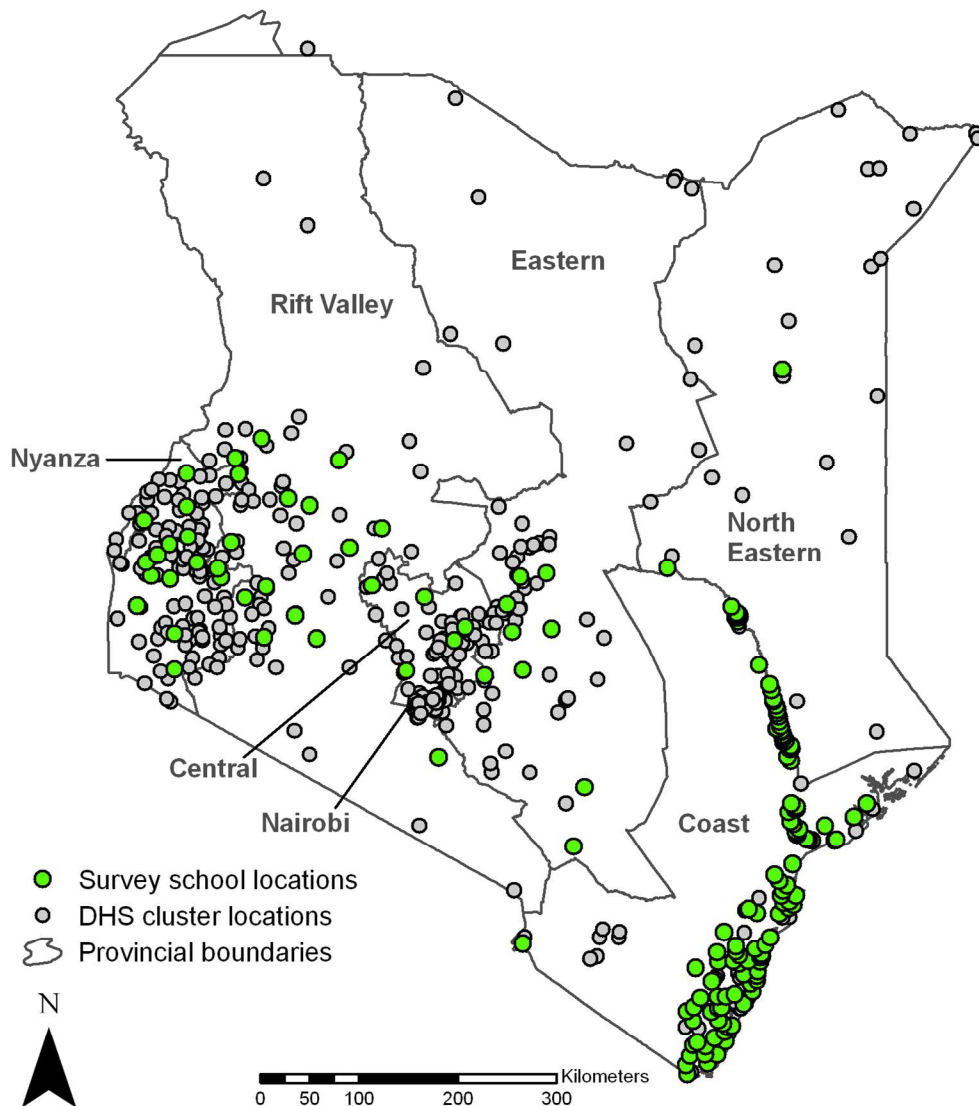
Proxy data for macro-nutritional status in school-aged children were derived from the Kenya 2008-09 Demographic and Health Survey (DHS). This dataset includes anthropometric measurements for 5,470 children aged under 6 years and 7,548 women of child-bearing age (ages 15 to 49). These were chosen to provide a general indication of the degree of malnutrition prevalent in the communities where school-children were surveyed. As DHS clusters rarely coincide with the location of surveyed schools (see Figure S1), it was first necessary to interpolate cluster-level prevalence to school locations using a weighted moving average approach (kriging). In order to provide a realistic account of uncertainty in both the predicted prevalence and the covariance functions used to define weighting, thus providing a range of potential prevalence values, this was embedded in a Bayesian framework using a simple binomial model structure.

The numbers of examined (n_i) and positive (Y_i) individuals for each anthropometric indicator in cluster i ($i=1, \dots, N$) were modelled as binomial variates in the form:

$$Y_i \sim \text{Binomial}(n_i, p_i)$$

$$\text{logit}(p_i) = \alpha + u_i$$

1 where α is the global mean prevalence and u_i a geostatistical random effect modelled using an isotropic,
 2 stationary exponential decay function: $f(d_{ab}; \phi) = \exp[-(\phi d_{ab})]$ where d_{ab} is the straight-line
 3 distance between cluster locations a and b, and ϕ is the rate of decline of spatial correlation. A non-
 4 distance between cluster locations a and b, and ϕ is the rate of decline of spatial correlation. A non-
 5 informative normal prior was used for α , the prior distribution of ϕ was uniform with upper and lower
 6 bounds set at 0.05 and 50 and the precision of u_i was given non-informative gamma distribution.
 7
 8 Following a burn-in of 10,000 iterations, the model was run for a further 10,000 iterations with thinning
 9 every ten iterations, during which predictions were stored.
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60



1 **Figure S1: Location of 167 school surveys and 400 DHS cluster sites across Kenya.** School survey data was
 2
 3 collected between 2008-2010, and DHS data between 2008-2009. Provincial boundaries are shown for
 4
 5 reference.
 6
 7
 8
 9
 10

11 **Table S1:** Summary of results from Bayesian kriging of anthropometric indicator data collected during 2008-
 12 2009 Kenya DHS survey.
 13

Anthropometric Indicator	Province	Mean observed prevalence at DHS clusters, % (range)	Mean predicted prevalence at school survey sites, % (range)
Wasting (in children < 5 yrs)	Central	3.3 (0, 25.0)	4.1 (3.9, 4.7)
	Coast	8.5 (0, 35.3)	10.5 (4.8, 22.8)
	Eastern	7.2 (0, 35.3)	5.1 (4.1, 6.4)
	North Eastern	12.6 (0, 33.3)	19.0 (12.7, 22.1)
	Nyanza	2.1 (0, 11.8)	2.5 (2.0, 3.7)
	Rift Valley	4.9 (0, 21.1)	4.1 (2.5, 9.1)
	Western	2.0 (0, 23.1)	2.3 (1.9, 3.6)
Low BMI (in women 15-49 yrs)	Central	8.5 (0, 4.0)	9.8 (5.6, 19.0)
	Coast	11.7 (0, 44.4)	16.8 (7.2, 28.5)
	Eastern	16.8 (0, 47.6)	13.6 (5.5, 23.9)
	North Eastern	27.6 (6.2, 52.9)	21.0 (16.3, 31.4)
	Nyanza	7.4 (0, 31.3)	7.2 (4.6, 9.7)
	Rift Valley	14.8 (0, 55.0)	11.8 (5.6, 22.3)
	Western	7.4 (0, 21.7)	8.2 (7.1, 9.3)
Consumed iron rich foods in past 24 hours (children < 5 yrs)	Central	14.8 (0, 45.5)	17.7 (9.1, 25.0)
	Coast	16.2 (0, 55.0)	17.5 (6.8, 31.4)
	Eastern	16.3 (0, 50.0)	18.1 (16.0, 20.7)
	North Eastern	15.3 (0, 60.7)	9.3 (5.9, 15.6)
	Nyanza	19.2 (0, 55.6)	21.4 (13.9, 27.7)
	Rift Valley	17.3 (0, 55.6)	17.2 (7.5, 23.8)
	Western	31.7 (0, 69.2)	28.3 (24.2, 35.3)

14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60

As can be seen from Table S1, observed prevalence of each indicator in the DHS dataset varied considerably within each Province. This is reflected in part by the range of predicted values shown for the school survey sites, although as geostatistical models tend to regress towards the mean the range in predicted values is considerably less. As under- and over-prediction are unlikely to be biased, this suggests that true association between proxies of malnutrition and hb levels in school-aged children are likely to be underestimated.

Bayesian hierarchical model of haemoglobin

For this model, haemoglobin level for each child i ($i = 1, \dots, N$) in school j ($j = 1, \dots, N$) is considered a normal variable outcome Y_{ij} in the form:

$$Y_{i,j} \sim \text{Normal}(\mu_{i,j}, \sigma^2)$$

$$\mu_{i,j} = \alpha + \sum_{k=1}^K \beta_k X_{i,j,k} + \sum_{l=1}^L \beta_l X_j + \beta_f u_j + \beta_g v_j + w_j$$

where α is the intercept, $\sum_{k=1}^K \beta_k X_{i,t,k}$ the matrix of individual-level covariates, $\sum_{l=1}^L \beta_l X_j$ the matrix of school-level covariates, u_j is considered a dummy variable labelled $u_j=1$ when prevalence of wasting is predicted to be medium/high severity (5-15%) and v_j a dummy variable labelled $v_j=1$ when prevalence of wasting is predicted to be extremely high (>15%) and w_j is a geostatistical random effect modelled using an isotropic, stationary exponential decay function: $f(d_{ab}; \phi) = \exp[-(\phi d_{ab})]$ where d_{ab} is the straight-line distance between pairs of school locations a and b , and ϕ is the rate of decline of spatial correlation. Non-informative priors were used for α and the coefficients (normal prior with mean 0 and precision 1 x 106), the prior distribution of ϕ was uniform with upper and lower bounds set at 1 and 50 and the precision of v_j was given a non-informative gamma distribution. Values of u_j and v_j at each model iteration are defined by a random variable Z_j , which is given an informative beta prior defined using the posterior mean and standard deviation of predicted prevalence of wasting (described above):

$$Z_j \sim \text{Beta}(\alpha_j, \beta_j)$$

$$0.05 > Z_j > 0.15, \quad u_j=1$$

$$0.15 > Z_j > 0.15, \quad v_j=1$$

Following a burn-in of 9,000 iterations, the values for the intercept and coefficients were stored for 1,000 iterations and model convergence was assessed using diagnostic tests for convergence and by visually inspecting the time series plots. Convergence was successfully achieved after 10,000, and the model was

run for a further 10,000 iterations with thinning every ten iterations, during which prediction locations were stored.

Model output therefore consisted of samples from the posterior distribution hb levels for individuals and summarised by school (mean Hb).

Model adjusted Population Attributable Risk Fractions (PAFs)

In general, analysis of adjusted PAFs for binary outcomes such as anaemia rely on using the adjusted odds ratios (ORs) resulting from binomial GLMs. Whilst regional PAFs can be estimated using this approach based on the spatial distribution of each risk factor by region, it cannot take into account spatial variation in the underlying haemoglobin levels of resident children and thus assumes the impact of each risk factor is constant. For this reason, we chose to use a two step process that explicitly considers variation in underlying haemoglobin levels to estimate PAFs for anaemia by Province across Kenya, as outlined below.

At each realisation of the Bayesian hierarchical model for haemoglobin, the mean hb level (μ_j) was estimated at each school for a baseline group (7 year old boys and 14 year old girls, in socio-economic quantiles 2-4) and for each risk group included in Table 4 (main text), including school-level covariates (SF status) and the geostatistical random error term (w_j). The normal cumulative distribution function (CDF) was then used to estimate the prevalence of anaemia (i.e. the proportion of children with hb levels lower than age- and sex-standardised cut-offs) for baseline and risk groups based on predicted μ_j and the modelled within-school standard deviation σ^2 , and the adjusted relative risk (RR) for each risk factor estimated. In order to produce internally valid estimates when using adjusted RRs, the method proposed by Kelinbaum *et al*⁽⁵²⁾ (see⁽³¹⁾ for discussion) was used to estimate the adjusted PAFs:

$$\text{PAF} = \sum_{i=0}^k pd_i \left(\frac{RR_i - 1}{RR_i} \right)$$

where pd_i is the proportion of cases falling into the i th exposure level, and is calculated as the observed school-level prevalence in either boys aged 6-8 years or girls aged 13-15, and RR_i is the relative risk the i th

1 exposure level with the unexposed group. School-level PAFs were then aggregated and summarised by
 2 Province. Model output therefore consisted of samples from the posterior distribution of Province mean
 3 PAFs for each risk factor.
 4
 5
 6
 7
 8
 9

10 **Model Validation**

11
 12 To assess the predictive performance of the final model of Hb, and associated predicted prevalence of Hb,
 13 hold-out validation data subsets were generated by random division of the data into quintiles. The
 14 Bayesian geostatistical model was then trained on four subsets whilst simultaneously predicting
 15 haemoglobin for the fifth, excluded subset. This was repeated for each subset, giving and observed and
 16 predicted Hb value for all surveyed individuals. The ability of the model to predict known mean Hb (and
 17 associated prevalence of anaemia) was assessed by three validation statistics: mean prediction error
 18 (which provides a measure of the bias of the predictor); absolute mean prediction error (which provides a
 19 measure of the accuracy of the predictor); and the correlation coefficient (providing a measure of
 20 association between observed and predicted Hb values).
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34

35 Results of this model validation process are shown below in Table S2. When summarised by school, there
 36 was a very high degree of correlation between observed and expected mean Hb. Mean error and absolute
 37 mean error of the anaemia predictions was -6%, suggesting that the model consistently over-predicted
 38 prevalence. At an individual level the model was less successful, on average producing a 10 g/L
 39 discrepancy between predicted and observed Hb levels. This difference between school and individual
 40 level predictive validity is reflective of the large residual variation between individuals within schools
 41 indicated by the model (Table 2).
 42
 43
 44
 45
 46
 47
 48
 49
 50

51 **Table S2:** Measures of correlation, bias and accuracy of Bayesian hierarchical regression models at individual and school
 52 levels.

53 Outcome	54 Mean error	55 Absolute mean error	56 Correlation coefficient
57 Individual-level:			
58 Hb level, g/L	- 0.005	+ 10.15	0.392
59 Anaemia	+ 0.042	+ 0.364	0.298
60 School-level:			
Mean Hb, g/L	- 1.74	+1.79	0.982

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Prevalence of anaemia, %	- 0.061	+0.065	0.951
--------------------------	---------	--------	-------

For peer review only

1
2
3
4
5
6
7 **Estimating the relative contribution of parasitic infections and nutrition for anaemia**
8
9 **among school-aged children in Kenya: a subnational geostatistical analysis**
10

11 Rachel L Pullan¹, Carol Gitonga^{1,2}, Charles Mwandawiro³, Robert W Snow^{2,4} and Simon J Brooker^{1,2}
12

13
14 ¹ *London School of Hygiene and Tropical Medicine, London, UK*

15 ² *Kenya Medical Research Institute-Wellcome Trust Research Programme, Nairobi, Kenya;*

16 ³ *Eastern and Southern Africa Centre of International Parasite Control, Kenya Medical Research Institute*
17 *(KEMRI), Nairobi, Kenya*

18 ⁴ *Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK*
19

20 **Correspondence:** rachel.pullan@lshtm.ac.uk; Faculty of Infectious and Tropical Diseases, London School
21
22 of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objectives: To quantify geographical variation in the relative contribution of parasitic infections, socio-economic factors and malnutrition in the aetiology of anaemia among school children across Kenya, thereby providing a rational basis for the targeting of an integrated school health package.

Design: Nationally representative cross-sectional survey data were collected using standard protocols. For all included children, data were recorded on haemoglobin concentration and common parasitic infections (*Plasmodium falciparum*, hookworm and schistosomiasis) and socio-economic indicators. Ecological proxies of malnutrition and food security were generated using Demographic and Health Survey and UN Food and Agriculture Organization food security data respectively. Spatially explicit, multilevel models were used to quantify impact upon child haemoglobin concentration.

Setting: Randomly selected schools in ecologically diverse settings across Kenya.

Main outcome measures: Mean haemoglobin concentration adjusted for infection, nutritional and socio-economic risk factors; associated risk ratios and adjusted Population Attributable Fractions (PAFs) for anaemia, by region.

Results: Data were available for 16,941 children in 167 schools; mean Hb was 122.1 g/L and 35.3% of children were anaemic. In multivariate analysis, mean Hb was significantly lower in boys and younger children. Severe malnutrition and interactions between *P. falciparum* and hookworm infections were significantly associated with lower Hb, with greater impacts seen for co-infected children. The contribution of risk factors to anaemia risk varied by province: in 14 year old girls PAFs ranged between 0-27.6% for *P. falciparum*, 0-29.0% for hookworm and 0-18.4% for severe malnutrition.

Conclusions: The observed geographical heterogeneity in the burden of anaemia attributable to different aetiological factors has important implications for the rational targeting of anti-anaemia interventions that can be included into an integrated school health programme.

Article Summary

Article Focus

- Cross-sectional haematological and parasitological school survey data for the period 2008-2010 in Kenya
- Spatially-explicit, Bayesian multi-level models quantifying impact of infection, nutrition and social risk factors on child haemoglobin concentration and adjusted Population Attributable Fractions for anaemia by province.
- Implications of spatial heterogeneity in underlying haemoglobin status and risk factors for cost-effective geographical targeting of [school-based anti-anaemia interventions, including deworming, malaria control and school feeding.](#)

Key messages

For school-aged children living in Kenya:

- *P. falciparum*, hookworm and schistosomes are all significantly associated with lower mean Hb, with the largest impact seen for coinfection with hookworm and *P. falciparum*, emphasising the importance of parasitic infections in the aetiology of anaemia in African school-age populations
- Poor food intake is suggested to be an important contributor to anaemia in school-aged children.
- There is pronounced geographical heterogeneity in the burden of anaemia attributable to these identified risk factors underscoring importance of appropriately targeting interventions

Strengths and limitations of the study

- Large and uniquely detailed dataset coupled with novel Bayesian modelling framework that allowed for simultaneous quantification of the mean shift in Hb associated with each risk factor, and sub-national risk of anaemia in school-aged children attributable to risk factors.
- Dependence upon cross-sectional design and ecological data limits inference of causality; incorporation of individual, nutritional-related factors and iron status indicators would improve quantification of their involvement.

Background

Children of school-going age (5-14 years) make up one quarter of the population of sub-Saharan Africa,¹ and the proportion of these children attending school has never been higher, in large part due to recent global and national commitments to universal primary education school enrolment.² Yet many children enrolled in school suffer from poor health and malnutrition and as a result are less likely to attend and complete school or achieve their full educational potential.³ For example, up to 40% of 5-14 year olds are estimated to be anaemic,^{4,5} which can affect cognition and learning.^{6,7} Anaemia also affects the health of older girls as they reach puberty, potentially impacting upon future motherhood and the nutrition of the next generation of learners.^{8,9} Fortunately, a number of simple and cost-effective interventions are available to avert the burden of anaemia, including malaria control,^{10,11} deworming,¹² iron supplementation^{13,14} and school-feeding¹⁵ – interventions which are increasingly delivered as part of an integrated school health package.¹⁶

Identifying the optimal combination of school health interventions requires reliable information on both the geographical distribution of anaemia and the relative contribution of different aetiological factors. A previous study of anaemia among preschool children in West Africa employed a geostatistical framework to interpolate estimates of anaemia among children aged 1-4 years collected through demographic health surveys (DHS), whilst incorporating ecological predictions of malaria and helminth infection. Whilst this work helped define the geographical distribution and limits of anaemia in the region it did not quantify spatial variation in the risk attributable to different aetiological factors.¹⁷ Here we take advantage of uniquely detailed haematological data collected during a comprehensive, national school health survey conducted across Kenya between 2008 and 2010.¹⁸ An initial analysis showed that the relative contribution of *Plasmodium falciparum* [to anaemia](#) varied according to malaria transmission zones.¹⁹ We extend this to investigate the relationships between haemoglobin concentration (Hb) and common parasitic infections (including *P. falciparum*, hookworm and schistosomes), socio-economic indicators and proxies of malnutrition and food security for school children across Kenya. We use an innovative Bayesian approach that quantifies the mean shift in Hb associated with each risk factor, which we use to

1
2
3
4
5
6
7 estimate sub-national variation in the risk of anaemia in school-aged children attributable to malnutrition,
8 malaria and helminths to define targets for integrated packages of school-health interventions.
9

10 11 12 13 **Methods**

14 15 **School survey data:**

16
17 This study uses cross-sectional school survey data collected between September 2008 and March 2010
18 using standard protocols, described elsewhere.^{18,20} In brief, 480 government primary schools were
19 selected to provide a national assessment of malaria and anaemia among school children. In a random
20 subset of 167 schools, stool and urine samples were also collected to assess prevalence and intensity of
21 helminth infections. At each school, 10 children plus 1 reserve of each sex were randomly selected from
22 each of classes 2-6 from the children present that day. [This captures children aged 4 to 16, although 80%](#)
23 [of included children were aged between 8 and 13 years.](#) Selected children were asked to provide a
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
fingerprick blood sample, which was used to assess haemoglobin (Hb) using a portable photometer
(Haemocue, Angelholm, Sweden) and *Plasmodium* infection in the peripheral blood based on microscopy-
corrected malaria rapid diagnostic test (RDT)¹⁸ results, [whereby slides were subsequently read for all RDT](#)
[positives.](#) Children in all schools were asked to provide stool samples which were examined for the eggs of
intestinal nematodes (*Ascaris lumbricoides*, *Trichuris trichiura* and hookworm species) and *S. mansoni*
using the Kato-Katz method. In schools on the coast, where *Schistosoma haematobium* is widespread,
children were asked to provide urine samples which were examined by urine filtration. A questionnaire
was administered to all pupils to obtain data on insecticide-treated net ownership and use, history of
anthelmintic medication, key socioeconomic variables, household construction and education of the
child's parent/guardian. School locations were determined using a hand-held Global Positioning System
and classified as urban using an updated 2010 urban extents (UE) mask derived from GRUMP.^{21,22}

50 51 **Nutritional proxy data**

1
2
3
4
5
6
7 Anthropometric and food intake data were not available for all of our study population, and we instead
8 relied on ecological measures. Proxy data for macro-nutritional status in school-aged children were
9 derived from the Kenya 2008-09 Demographic and Health Survey (DHS).^{23, 24} Anthropometric
10 measurements were extracted for children ≤5 years and for women of child-bearing age (ages 15 to 49)
11 (see Table 1). For children ≤5 years, the reported proportion consuming food rich in iron (flesh and organ
12 meat) in the past 24 hours was also extracted. These indicators, which provide a general measure of
13 nutritional status in the wider community, were geo-located to the centre of each sampling cluster.
14
15 Bayesian kriging²⁵ [a geostatistical interpolation method that accounts for the error introduced by](#)
16 [estimating the semivariogram model \(a function of the variability in outcome against the distance](#)
17 [separating observation points\)](#), was used to interpolate a distribution of possible values for each indicator
18 to survey school locations. [This was done by incorporating the full posterior distribution of the Bayesian](#)
19 [semivariogram model, estimated](#) -using the Bayesian statistical software WinBUGS version 14.1 (Medical
20 Research Council Biostatistics Unit and Imperial College London). Mean predicted prevalence values were
21 then categorised according to WHO-defined cut-off values for severity of malnutrition by prevalence
22 range (Table 1). See Technical Appendix for further details.
23
24
25
26
27
28
29
30
31
32
33
34

35 Two additional indicators were used to describe general food availability at each school location:
36 predominant source of livelihood (based on 2010 FAO Livelihood zones^{26, 27}); and food security conditions
37 at the time of the survey (based on the FAO Food Insecurity Severity Scale^{27, 28}). Both indicators were
38 obtained from the Famine Early Warning System Network (FEWS NET; <http://www.fews.net>). Information
39 on the intended provision of formal, government-run school-feeding programmes was abstracted from
40 the Homegrown School Meals Programme (HGSM) and the Emergency HGSM 2010 registers provided
41 by the Kenya Ministry of Education (*personal communication*).
42
43
44
45
46
47
48

49 **Analysis**

50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 A schematic summarising the major steps in our analysis strategy is provided in Figure 1. Haemoglobin
8 levels were adjusted for altitude and individuals were classified as anaemic using age and sex specific cut-
9 offs.²⁹ Children were classified as having medium/high intensity hookworm infection if infection intensity
10 exceeded 100 epg, and medium/high intensity *S. haematobium* infection if infection intensity exceeded
11 500 epg (both representing the top 10th percentile). Reported information on ownership of household
12 assets and household construction was used to generate wealth indices using principle component
13 analysis,³⁰ and resulting scores were divided into quintiles.

14
15
16
17
18
19
20
21 The initial candidate set of predictor variables for univariate analysis included age, sex, relative socio-
22 economic status, education-level of the household head, protective behaviours (bednet usage and
23 deworming) and infection status with hookworm, schistosome and *P. falciparum*, in addition to school-
24 level ecological covariates including urban/rural location, predicted prevalence of anthropometric
25 indicators in younger children and women of reproductive age, livelihood zone and food security status
26 and intended provision of school-feeding. In order to select candidate variables for spatial multivariable
27 analysis, univariate analysis was undertaken using logistic regression (criteria: Wald test $P < .1$) using Stata
28 11.1 software (StataCorp). Standard errors were adjusted for dependence between individuals within
29 schools using a clustered sandwich estimator.³¹ Backward-stepwise elimination was then used to generate
30 a minimum adequate multivariate model; excluded covariates ($P > .05$) were retested in the minimal
31 model to confirm lack of association. [Basic age-sex interaction terms were also investigated, but there](#)
32 [was insufficient evidence that these improved the overall fit of the model.](#)
33
34
35
36
37
38
39
40
41
42
43

44 Retained covariates were included in a spatial multivariable Bayesian mixed-effects linear regression
45 model using WinBUGs 14.1. This model included individual and school-level ecological fixed effects and a
46 school-level geostatistical random effect (using an isotropic, stationary exponential decay function),²⁵
47 thereby incorporating a degree of spatial smoothing. Uncertainty in the predicted prevalence of
48 anthropometric proxy indicators was incorporated by modelling these covariates as random variables
49
50
51
52
53
54

1
2
3
4
5
6
7 with a beta distribution, defined using the prediction posterior distribution. At each model realisation,
8 malnutrition was categorised as either low, medium/high or severe for each school location, according to
9 Table 1.
10

11
12
13
14 Given the large observed differences in mean Hb and prevalence of predictors by region and demographic
15 group, adjusted [Population Attributable Fractions \(PAFs\)](#) for anaemia were estimated for two indicator
16 demographic groups (boys aged 7 years, and girls aged 14 years) for each province in Kenya. In brief, at
17 each model realisation we estimated the prevalence of anaemia in each of the risk groups based upon the
18 posterior mean and standard deviation of the Hb distribution within each school. These were used to
19 estimate (i) the relative risk of anaemia for each risk group compared with baseline risk, and (ii) the
20 associated adjusted PAF, as described by Rockhill *et al* (1998).³² Estimates were subsequently
21 summarised by province. Full statistical notation of Bayesian geostatistical modelling, PAF and model
22 validation procedures are presented in Technical Appendix. Model validation using hold out datasets,
23 presented in detail in the accompanying Technical Appendix, suggest very high degree of correlation at
24 the school level between observed and expected mean Hb levels ($R^2 > 0.98$), and prevalence of anaemia
25 ($R^2 > 0.95$) suggesting the approach we used to estimate anaemia prevalence (and associated attributable
26 fractions) was appropriate.
27
28
29
30
31
32
33
34
35
36
37
38
39
40

41 **Results:**

42
43 Data on Hb, malaria parasitaemia and helminth infection were available for 16,941 children (aged 5-16) in
44 167 schools (Figure 2). Mean altitude-adjusted Hb concentration was 122.1 g/L, with 35.3% of children
45 estimated to be anaemic. The spatial distribution of the unadjusted Hb data is presented in Figure 2,
46 summarised by school. There were significant geographical heterogeneity in mean Hb across the country
47 (Moran's I $p < 0.001$), with the lowest average levels observed in Coast Province (120.1 g/L) and the highest
48 in Central Province (134.0 g/L). Univariate analysis suggested that Hb levels increased with age and with
49
50
51
52
53
54

1
2
3
4
5
6
7 relative socio-economic status, and were significantly lower in individuals infected with *P. falciparum* or
8 medium/high intensity helminth infections (Table 1). Hb was also significantly lower in areas with
9 predicted severe malnutrition, although there was no ecological relationship with predicted consumption
10 of iron rich foods. Univariate analysis also suggests a relationship between Hb and food availability, with
11 significantly lower Hb levels in children living in very food insecure areas.
12
13
14

15 16 17 18 **Bayesian hierarchical model of haemoglobin**

19
20 Table 3 presents a multivariate analysis of factors associated with Hb levels. On average, boys had lower
21 mean Hb compared with girls by 1.53 g/L (95% BCI: 1.13-1.94 g/L), and older children tended to have
22 higher Hb levels, increasing 1.08 g/L (95% BCI: 0.99-1.17 g/L) each yearly age interval. There was strong
23 evidence for a multiplicative negative interaction between infection status and Hb levels, highlighting the
24 health impacts of parasitic infections in school-aged children. Children with medium/high intensity
25 hookworm infection had Hb levels 1.03 g/L (95% BCI: 0.02-2.00 g/L) lower than uninfected children, those
26 with *P. falciparum* malaria 3.00 g/L (95% BCI: 1.68-4.29 g/L) lower and *S. haematobium* infection 1.75 g/L
27 (95% BCI: 0.60-2.92 g/L); children co-infected with both *P. falciparum* malaria and medium/high intensity
28 hookworm infection had Hb levels 5.56 g/L (95% BCI: 2.67-8.64 g/L) lower than uninfected children.
29
30
31
32
33
34
35

36
37
38 Proxy measures of nutritional status were also strongly associated with Hb levels. School-children living in
39 areas predicted to have severe prevalence of wasting had Hb levels that were on average 9.74 g/L (95%
40 BCI: 7.45-12.16 g/L) lower than those living in low risk areas. School-feeding programmes also appear to
41 have been targeted to districts with increased risk of anaemia, although examination of the 95% credible
42 intervals suggests no statistical difference within target districts between mean Hb levels in schools with
43 and without intended provision of formal school-feeding programmes. Perhaps surprisingly, food security
44 at the time of the survey (a qualitative measure that allows adjustment for acute conditions and the
45 impact of seasonality) was not associated with Hb levels in multivariate analysis. Similarly, after
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 accounting for all other factors in the model, relative socio-economic status was not significantly
8 associated with Hb levels.
9

10
11
12 The random effects variance components from this model indicate that most of the residual variation
13 remains between children within schools: of the total unexplained variability in Hb, 13.8% occurred at the
14 school level, whilst 86.2% occurred between children within schools. Closer examination of the spatial
15 school-level random effect suggests that, after accounting for included predictors, most schools (81%) had
16 similar Hb levels that could not be distinguished statistically (shown in green in Figure 23). There was
17 however some residual spatial clustering of schools with higher/lower than expected Hb levels (foci radii
18 of ~190 km, Table 2) with significantly lower mean Hb levels observed in schools in the far south of Coast
19 Province and in Nyanza and Western provinces (Figure 3). Interestingly, comparison of observed
20 prevalence of *P. falciparum* and medium/high intensity hookworm suggests that prevalence of both
21 infections was higher for these schools than for other schools in these provinces (*P. falciparum* Kruskal-
22 Wallis p 0.02; medium/high intensity hookworm Kruskal-Wallis p 0.01). In contrast, mean Hb levels were
23 significantly higher for schools in Rift Valley, Central and Eastern provinces
24
25
26
27
28
29
30
31
32
33
34
35

36 **Adjusted Population Attributable Risk Fractions**

37
38 Using the final Bayesian hierarchical model, we estimated for two demographic groups (girls aged 14
39 years, and boys aged 7 years) the PAF of anaemia due to *P. falciparum*, medium/high intensity hookworm
40 infection, medium/high intensity *S. haematobium* infection, and acute malnutrition (prevalence of
41 wasting in under 5s exceeding 15%) for each province in Kenya, adjusting for all other included predictors
42 (Table 4). Our results suggest that the relative contribution of each risk factor varies considerably by
43 province. For example, in Coast Province 13% (95% BCI: 4-23%) of anaemia in 14 year old girls attending
44 school may be attributable to medium/high intensity hookworm, comparing with 6% (95% BCI: 2-10%) to
45 *P. falciparum* infection and 3.1% (95% BCI: 0-8%) to schistosomiasis. Meanwhile, results suggest that
46 reduction in acute malnutrition in Coast Province (to <5% wasting in ≤5s) would reduce anaemia
47
48
49
50
51
52
53
54

1
2
3
4
5
6
7 prevalence by 19% (95% BCI: 0-38%). This contrasts with Western Province, where the estimated risk of
8 anaemia attributable to medium and high intensity hookworm is 29% (95% BCI: 1-55%) and to *P.*
9 *falciparum* 24% (95% BCI 0-41%); no school children in this district are exposed to *S. haematobium* and
10 acute malnutrition in under 5s is predicted not to exceed 5%. The wide 95% BCIs observed reflect (i) the
11 smaller sample sizes for provinces other than Coast (given that N = schools, not individuals), and (ii) the
12 additional uncertainty surrounding proxies for malnutrition. Trends in PAFs observed for predictors in 7
13 year old boys mirror those for older girls but are generally lower, a reflection of the lower baseline Hb
14 levels in this group.
15
16
17
18
19
20
21
22
23
24

25 Discussion

26
27 As global commitment and resources for school health and nutrition initiatives grow,¹⁶ there is an urgent
28 need to strengthen our understanding of the likely impact of interventions aimed at tackling important
29 health and development outcomes such as anaemia. This requires a detailed understanding of the relative
30 contribution of aetiological factors at policy relevant scales. Here we present a uniquely detailed national
31 analysis of risk factors for anaemia among schoolchildren in 167 schools across Kenya. We introduce a
32 new approach using multilevel linear modelling to provide a detailed description the varying contribution
33 of risk factors by administrative unit (province), whilst fully accounting for spatial variation in underlying
34 mean Hb. Results provide a vital decision-making tool for cost-effective targeting the delivery of
35 integrated, packages of interventions tailored to each province, with the aim of reducing anaemia in
36 schoolchildren. [For example, school-feeding programmes in Coast, Eastern and Rift Valley pProvinces may
37 want to consider including iron-rich food stuffs, whilst programmes in Nyanza and Western pProvinces
38 omitting malaria control initiatives may not see desired improvements in childhood anaemia."](#)
39
40
41
42
43
44
45
46
47
48
49

50 Our analysis found evidence that *P. falciparum*, hookworm and schistosomes were each significantly
51 associated with lower mean Hb, observations consistent with current knowledge. Such findings
52
53
54

1
2
3
4
5
6
7 corroborate recent results from the Global Burden of Disease study, which suggests that globally for
8 anaemia 10.2% of years lived with a disability (YLD) across all ages could be attributed to hookworm,
9 17.9% to malaria and 3.4% to schistosomiasis.³³ Our analysis provides three additional insights into the
10 epidemiology of anaemia. First, we confirm small-scale studies³⁴⁻³⁸ that have suggested that coinfection
11 with hookworm and *P. falciparum* may have a multiplicative effect on Hb level – effects which occur
12 across Kenya. Second, we demonstrate pronounced geographical heterogeneity in the burden of anaemia
13 attributable to identified risk factors. Associations for hookworm and *P. falciparum* were most
14 pronounced in Western Province, where results suggest that together these two infections were
15 attributable for close to 53% (95% BCI: 1.4-96.0%) of anaemia cases in girls aged 14 years. Together, these
16 findings emphasise the importance of parasitic infections in the aetiology of anaemia in school-aged
17 children and underscore current efforts to control helminths and malaria as part of integrated school
18 health programmes.^{39, 40} The PAFs provided here are useful for translating our findings into numbers that
19 can help policy decision makers appreciate the potential benefits of carefully targeted programmes
20 tailored to each province that strike the best balance between dietary interventions, deworming and
21 malaria control.^{17, 41}

22
23
24
25
26
27
28
29
30
31
32
33
34
35 Third, whilst acknowledging that we rely on proxy indicators, our results do suggest that poor food intake
36 is an important contributor to anaemia in school children: the final model indicates that in communities
37 where acute malnutrition is a severe public health problem (i.e. where wasting in under 5s exceeds 15%),
38 mean Hb levels in school-aged children are on average 9.7 g/L lower than well nourished communities,
39 and as such eliminating acute malnutrition as a public health problem could avert close to 20% of anaemia
40 cases in 14 year old girls. This is consistent with limited evidence from smaller individual-level studies
41 suggesting an association between under-nutrition and anaemia in school-aged populations.^{42, 43} Our
42 results also provide some indication that whilst existing school feeding programmes in Kenya do appear
43 to be targeted to those at most need, they are currently ineffective at tackling anaemia in this population.
44
45
46
47
48
49
50
51
52
53
54

1
2
3
4
5
6
7 Our spatially explicit hierarchical approach seeks to maximise the usefulness of this observational, cross-
8 sectional data and results show that whilst much of the between school variation could be explained by
9 included risk factors, a high degree of variation does still remain within schools. This suggests that there
10 are individual-level factors beyond those measured that still need to be considered. In addition to
11 individual iron status or food intake there are a number of candidates. HIV-infection is potential likely to
12 be a contributor, as even in asymptomatic HIV infected individuals anaemia is a common manifestation.⁴⁴
13
14
15
16
17 ⁴⁵ We were also unable to account for genetic traits such as sickle-cell and other haemoglobinopathies,⁴⁶
18 or for the potential effect of menarche in adolescent girls.⁴⁷
19
20
21
22

23 A number of additional study limitations should also be acknowledged. As we have previously
24 acknowledged,²⁰ cross-sectional studies such as this are inevitably subject to a number of limitations
25 including the potential of ecological fallacy, especially when incorporating population-level estimates. We
26 acknowledge that our indicator of school-feeding (inclusion in the government school-feeding registers) is
27 far from optimal, and the study would have benefitted from detailed individual-level data on iron status
28 and food intake. In addition, modelled associations are likely to be lower than the true impact, due to
29 non-systematic measurement error in our infection indicators, which rely on single samples using
30 standard parasitological diagnosis with <100% sensitivity.^{48,49} Despite a robust stratified random sampling
31 frame, our results may also be subject to some selection bias, as very ill children such as those with acute
32 infections and severe chronic malnutrition may have been absent on the day of the school survey and
33 thus not included in the study sample. [Although contemporary regional or school-level attendance data is
34 not available, overall school attendance rates are known to vary geographically across Kenya.⁵⁰ For
35 example, in 2005 \(the most recent year for which sub-national net primary school attendance figures are
36 available\) reported net attendance in Nyando District \(Nyanza Province\) was greater than 90% for both
37 males and females, whilst in Turkana District \(Rift Valley Province\) levels were between 30-40%.⁵⁰ If Hb
38 levels differ systematically between non-attending and attending school-aged children, this may act to
39 either dilute or exaggerate geographical variation in Hb levels](#)
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54

1
2
3
4
5
6
7 ~~Neither is the~~ [An additional factor that would influence individual-level variation is the onset of menarche](#)
8 [in girls: no specific data were collected to record this, and inclusion of a basic age-sex interaction term did](#)
9 [not improve model fit suggesting that more detailed information is required. Finally, this](#) analysis is unable
10 to account for the impact of recently cleared infections on current Hb status. ~~Finally, the current and the~~
11 [presented](#) results represent a snapshot of the period 2008-2010 and it is conceivable that seasonality and
12 intervention-related factors may have confounded results. However, the inclusion of mosquito net use,
13 recent deworming or food security (a seasonal covariate) in multivariate models did not improve model
14 fits, suggesting that temporal and control factors were relatively unimportant.

21 22 23 **Conclusions**

24
25 A recent focus on school-child and adolescent health has emphasised the need to develop a deeper
26 understanding of factors impacting health and development of this population group.⁵¹⁻⁵⁴ Here we
27 present a novel two-step process to model Hb levels and then predict associated anaemia prevalence in
28 different areas of Kenya. Our approach adds considerable value over and above what could be achieved
29 modelling anaemia directly, and the adopted geostatistical Bayesian framework allows for propagation of
30 uncertainty associated at each step (including generation of predicted ecological covariates) into the final
31 confidence limits. Future expansion of this model could look at interpolating existing results to make
32 predictions in non-surveyed areas, although this would require generation of values for all included
33 individual-level covariates for each prediction location, increasing prediction uncertainty. Our results
34 highlight the marked geographical heterogeneity in the burden of anaemia attributable to identified risk
35 factors in Kenya, which has major implications for effective targeting of packages of school-based
36 interventions. [We suggest that deworming can help tackle anaemia both in western and coastal Kenya,](#)
37 [whereas malaria control is likely to have the greatest impact in western Kenya, whilst school-feeding](#)
38 [programmes can help reduce anaemia in Coast, Eastern and Rift Valley provinces.](#) The priority now should
39 be to explore in detail the health and educational impact of carefully designed integrated infection and
40 nutritional interventions in this age group.
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Acknowledgements

We are very grateful to all the teachers, parents and especially children who participated in this survey and to the many facilitators and members of the survey teams who collected the data [and to Don Bundy for comments](#). The survey work was supported by the Wellcome Trust UK as part of fellowship support to SJB (#81673) and by the Division of Malaria Control, Ministry of Public Health and Sanitation through a grant from the Department for International Development (DfID) through the WHO Kenya Country Office to Dr AM Noor and Dr E Juma.

Footnotes

Contributors RLP was responsible for designing and conducting analysis, interpretation, drafting and production of the final manuscript. CG contributed to the survey design, data assembly and cleaning and contributed to final manuscript. CM and RWS contributed to survey design, interpretation of results and contributed to final manuscript. SJB [designed the field surveys](#), provided scientific guidance and contributed to interpretation and preparation of the final manuscript. All authors read and approved the final manuscript.

Funding The fieldwork was supported by the Wellcome Trust UK and by the Division of Malaria Control, Ministry of Public Health and Sanitation. RLP acknowledges support of the Bill & Melinda Gates Foundation and CWG is supported by a Commonwealth Scholarship from the DfID. SJB is supported by a Wellcome Trust Senior Fellowship in Basic Biomedical Science (098045) and RWS is supported by a Wellcome Trust Principal Fellowship (079080).

Competing interests None.

Data sharing statement: The school survey dataset used for this analysis is available for download through the Global Atlas of Helminth Infection (www.thiswormyworld.org). The WinBUGs statistical code is available from the corresponding author at rachel.pullan@lshtm.ac.uk.

References

1. United Nations Department of Economic and Social Affairs. World population prospects, 2010 revision; Available from: <http://esa.Un.Org/unpd/wpp/>.
2. Global Partnership for Education. All children learning: Report 2011. Global Partnership for Education; Available from: <http://allchildrenlearningreport.globalpartnership.org>
3. World Bank. Fresh: Focusing resources on education and school health. A fresh start to improving the quality and equity of education. UNESCO, UNICEF, WHO & World Bank; 2000
4. McLean E, Cogswell M, Egli I, Wojdyla D, de Benoist B. Worldwide prevalence of anaemia, WHO vitamin and mineral nutritional information system, 1993-2005. *Public Health Nutrition*. 2008;12(4):444-454.
5. Best C, Neufingerl N, van Geel L, van den Briel T, Osendarp S. The nutritional status of school-aged children: Why should we care? *Food and Nutrition Bulletin*. 2010;31(3):400-417.
6. Grantham-McGregor S, Ani C. A review of studies on the effect of iron deficiency on cognitive development in children. *Journal of Nutrition*. 2001 Feb;131(2S-2):649S-666S.
7. Fretham SJ, Carlson ES, Georgieff MK. The role of iron in learning and memory. *Advances in Nutrition*. 2011;2(2):112-121.
8. Scholl TO. Iron status during pregnancy: Setting the stage for mother and infant. *American Journal of Clinical Nutrition*. 2005;81(5):1218S-1222S.
9. Dharmalingam A, Navaneetham K, Krishnakumar CS. Nutritional status of mothers and low birth weight in india. *Maternal and Child Health Journal*. 2010;14(2):290-298.
10. Clarke SE, Jukes M, Njagi JK, Khasakhala L, Cundill B, Otido J, et al. Effect of intermittent preventive treatment of malaria on health and education in schoolchildren: A cluster-randomised, double-blind, placebo-controlled trial. *Lancet*. 2008;372(9633):127-138.
11. Leenstra T, Phillips-Howard PA, Kariuki SK, Hawley WA, Alaii JA, Rosen DH, et al. Permethrin-treated bed nets in the prevention of malaria and anemia in adolescent schoolgirls in western kenya. *American Journal of Tropical Medicine and Hygiene*. 2003;68(S86-S93).
12. Smith JL, Brooker S. The impact of hookworm infection and deworming on anaemia in non-pregnant populations: A quantitative update. *Tropical Medicine and International Health*. 2010;15(7):776-795.
13. Leenstra T, Kariuki SK, Kurtis JD, Oloo AJ, Kager PA, ter Kuile FO. The effect of weekly iron and vitamin a supplementation on hemoglobin levels and iron status in adolescent schoolgirls in western kenya. *European Journal of Clinical Nutrition*. 2009;63(2):173-182.
14. Gera T, Sachdev HP, Nestel P. Effects of iron supplementation on physical performance in children and adolescents: Systematic review of randomized controlled trials. *Indian Pediatrics*. 2007;44:15-24.
15. Bundy DAP, Burbano C, Grosh M, Gelli A, Jukes M, Drake L. *Rethinking school feeding: Social safety nets, child development and the education sector*. The World Bank: New York; 2009.
16. Bundy DAP. *Rethinking school health: A key component of education for all*. Washington: The World Bank; 2012.
17. Soares Magalhães RJ, Clements ACA. Mapping the risk of anaemia in preschool-age children: The contribution of malnutrition, malaria, and helminth infections in West Africa. *PLoS Medicine*. 2011;8(6):e1000438.
18. Gitonga C, Karanja P, Kihara J, Mwanje M, Juma E, Snow R, et al. Implementing school malaria surveys in kenya: Towards a national surveillance system. *Malaria Journal*. 2010;9(1):306.
19. Gitonga CW, Edwards T, Karanja PN, Noor AM, Snow RW, Brooker SJ. *Plasmodium* infection, anaemia and mosquito net use among school children across different settings in Kenya. *Tropical Medicine and International Health*. 2012;17(7):858-870.
20. Brooker S, Pullan RL, Gitonga CW, Ashton RA, Kolaczinski JH, Kabatereine NB, et al. *Plasmodium*-helminth coinfection and its sources of heterogeneity across east africa. *Journal of Infectious Diseases* 2012;205(5):841-852.
21. Balk DL, Deichmann U, Yetman G, Pozzi F, Hay SI, Nelson A. Determining global population distribution: Methods, applications and data. *Advances in Parasitology*. 2006;62:119-156.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
22. Balk D, Pozzi F, Yetman G, Deichmann U, Nelson A. The distribution of people and the dimension of place: Methodologies to improve the global estimation of urban extents. Urban Remote Sensing Conference; 2004; Available at: <http://www.iussp.org/Activities/wgc-urb/balk.pdf>
 23. Kenya National Bureau of Statistics, Measure DHS. Kenya: Dsh, 2008-09 - final report. Calverton, Maryland, USA: Measure DHS, ICF Macro; 2010
 24. Measure DHS. Kenya 2008-09 DHS data. Available from: <http://www.measuredhs.com>
 25. Diggle P, Moyeed R, Tawn J. Model-based geostatistics (with discussion). *Applied Statistics*. 1998;47:299-350.
 26. FEWSNET. Famine early warning system network (FEWSNET) livelihood zone map for Kenya. FEWSNET and USAID; 2011
 27. Grillo J, Holt J. Application of the livelihood zone maps and profiles for food security analysis and early warning: USAID FEWSNET; 2009
 28. FEWSNET. Famine early warning system network (FEWSNET) Kenya archive: Food security update. FEWSNET and USAID; 2011
 29. WHO. Iron deficiency anaemia: Assessment, prevention and control. A guide for programme managers. Geneva: World Health Organisation; 2001
 30. Filmer D, Pritchett LH. Estimating wealth effects without expenditure data--or tears: An application to educational enrollments in states of India. *Demography*. 2001;38(1):115-32.
 31. Williams RL. A Note on Robust Variance Estimation for Cluster-Correlated Data. *Biometrics*. 2000; 56: 645-6.
 32. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *American Journal of Public Health*. 1998;88(1):15-19.
 33. Jasrasaria J, Murray CJL, Naghavi M. Global Burden of Disease 2010 Estimate: Strategy Report for Anaemia Envelope. [Unpublished report]. *Personal Communication* 2012.
 34. Brooker S, Akhwale WS, Pullan R, Estambale B, Clarke S, Snow RW, et al. Epidemiology of *Plasmodium*-helminth coinfection in Africa: Potential impact on anaemia and prospects for combining control. *American Journal of Tropical Medicine & Hygiene*. 2007;77:88-98.
 35. Pullan R, Brooker S. The health impact of polyparasitism in humans: Are we under-estimating the burden of parasitic diseases? *Parasitology*. 2008;1-12.
 36. Nkwo-Akenji TK, Chi PC, Cho JF, Ndamukong KKJ, Sumbele I. Malaria and helminth co-infection in children living in a malaria endemic setting of Mount Cameroon and predictors of anaemia. *Journal of Parasitology*. 2006;92(6):1191-1195.
 37. Midzi N, Mtapuri-Zinyowera S, Mapingure MP, Sangweme D, Chirehwa MT, Brouwer KC, et al. Consequences of polyparasitism on anaemia among primary school children in Zimbabwe. *Acta Trop*. 2010;115:103-111.
 38. Ezeamama EE, McGarvey ST, Acosta LP, Zierler S, Manalo DL, Wu H-W, et al. The synergistic effect of concomitant schistosomiasis, hookworm, and trichurias infections on children's anemia burden. *PLoS Negl Trop Dis*. 2008;2(6):e245.
 39. Bundy DAP, Shaeffer S, Jukes M, Beegle K, Gillespie A, Drake L, et al. School based health and nutrition programs. *Chapter 58 in "Disease Control Priorities for Developing Countries"*. 2006.
 40. Brooker S, Clarke S, Snow RW, Bundy DAP. Malaria in african schoolchildren: Options for control. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2008;102(4):304-305.
 41. Rowe AK, Powell KE, Flanders WD. Why population attributable fractions can sum to more than one. *American Journal of Preventative Medicine*. 2004;26:243-249.
 42. Anumudu C, Adolami M, Igwe C, Nwagwu M, Keshinro O. Nutritional anaemia and malaria in pre-school and school-aged children. 2008;1(11-17).
 43. Daboné C, Delise HF, Receveur O. Poor nutritional status of schoolchildren in urban and peri-urban areas of Ouagadougou (Burkina Faso). *Nutrition Journal*. 2011;10:34.
 44. Calis JC, van Hensbroek MB, de Haan RJ, Moons P, Brabin BJ, Bates I. HIV-associated anemia in children: A systematic review from a global perspective. *AIDS*. 2008;22(10):1099-1112.
 45. Levine AM, Berhane K, Masri-Lavine L, Sanchez M, Young M, Augenbraun M, et al. Prevalence and correlates of anemia in a large cohort of HIV-infected women: Women's interagency HIV study. *Journal of Acquired Immune Deficiency Syndrome*. 2001;26:28-35.

- 1
2
3
4
5
6
7 46. Tolentino K, Friedman JF. An update on anaemia in less developed countries. *American Journal of Tropical Medicine and Hygiene*. 2007;77:44-51.
- 8 47. Leenstra T, Kakiuki SK, Kurtis JD, Oloo AJ, Kager PA, ter Kuile FO. Prevalence and severity of
9 anaemia and iron deficiency: Cross-sectional studies in adolescent schoolgirls in western Kenya. *European
10 Journal of Clinical Nutrition*. 2004;58(4):681-691.
- 11 48. Booth M, Vounatsou P, N'Goran E K, Tanner M, Utzinger J. The influence of sampling effort and
12 the performance of the kato-katz technique in diagnosing *Schistosoma mansoni* and hookworm co-
13 infections in rural Cote d'Ivoire. *Parasitology*. 2003;127(Pt 6):525-31.
- 14 49. Okell LC, Ghani AC, Lyons E, Drakeley C. Submicroscopic infection in *Plasmodium falciparum*
15 endemic populations: A systematic review and meta-analysis. *Journal of Infectious Disease*.
2009;15(10):1509-1517.
- 16 50. World Bank. World development report 2007: Development and the next generation.
17 Washington: The World Bank
- 18 51. Education and Policy Data Center. Kenya 2005 District Primary Education Profiles. Available From:
19 www.epdc.org/country/kenya
- 20 52. Little AW, Green A. Successful globalisation, education and sustainable development.
21 *International Journal of Education and Development*. 2009;29:166-174.
- 22 53. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S-J, Dick B, Ezech AC, et al. Adolescence: A
23 foundation for future health. *The Lancet*. 2012;379(9826):1630-1640.
- 24 54. Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiological research. Belmont, CA: Lifetime
25 Learning Publications; 1982. p. 163.
- 26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7 **Figure legends:**

8 **Figure 1 Schematic of adopted analysis strategy.** The dual aim of this analysis strategy was to (i) model
9 the impact of risk factors on children's haemoglobin (Hb) levels, and (ii) estimate the adjusted Population
10 Attributable Risk Fraction (PAF) associated with important, modifiable risk factors. This process was
11 repeated at each realisation of the Bayesian hierarchical model to provide posterior mean estimates and
12 credible intervals for all parameters of relevance.

13 **Figure 2 Spatial distribution of mean haemoglobin distribution across Kenya, by school.** HB is adjusted
14 for altitude, population density (persons / km²) is shown for reference.

15 **Figure 3 Graph of spatial random-effect (school-level residual) from the multivariate, Bayesian**
16 **hierarchical model for Hb.** Each point represents one school; the posterior mean is shown by the dot and
17 the 95% BCI are represented by the error bars. The horizontal straight line in the middle of the graph
18 represents the overall mean of the residuals, which are zero centred. Inset shows the school locations,
19 shaded by the value of the school-level residual.

1
2
3
4
5
6
7 **Tables:**
8

9 **Table 1** Definition of included anthropometric indicators, and population-based cut-off values for severity of
10 malnutrition, as defined by the World Health Organization.
11

Anthropometric Measurement		Severity of malnutrition (prevalence range for < 6 years)		
		Low severity	Medium/high severity	Very high severity
Stunting	Height-for-age Z-score > -2	<20%	20-40%	>40%
	SD			
Wasting	weight-for-height Z-score >	<5%	5-15%	>15%
	-2 SD			
Underweight	weight-for-age <i>or</i>	<10%	10-30%	>30%
	BMI-for-age Z-score > -2 SD			

12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 2 Mean Hb concentration according to infection status, nutritional proxies and other characteristics in 16,941 children aged 5-16 years (univariate analysis).

Variable	Data type and source	Children No (%)	Mean Hb shift, g/L (SE)	p
Sex				
Female	Individual – KHS	8,314 (49.1)	(vs. female)	
Male	Individual – KHS	8,627 (50.9)	- 1.36 (0.28)	<0.001
Age (in years)	Individual – KHS	-	+ 0.89 (0.11)	<0.001
Socio-economic indicators				
Educated household head ¹	Individual – KHS	8,333 (49.2)	+ 0.82 (0.41)	0.05
Household owns a mobile phone	Individual – KHS	10,065 (59.4)	+ 1.42 (0.40)	<0.001
SES quintile				
Lowest	Individual – KHS	3,272 (20.4)	(vs. lowest)	
Mid (quintiles 2-4)		9,926 (61.8)	+ 2.35 (0.50)	<0.001
Highest		2,854 (17.8)	+ 3.07 (0.82)	<0.001
Urban location	School-level – digital UE map	1,237 (7.3)	1.56 (1.52)	0.3
Current infections				
<i>Malaria-P. falciparum</i> parasitaemic	Individual – KHS	607 (3.6)	-3.36 (0.77)	<0.001
Med/high intensity hookworm (>100 epg)	Individual – KHS	931 (5.5)	-2.81 (0.69)	<0.001
Med/high intensity <i>S. haematobium</i> (>500 epg)	Individual – KHS	799 (4.7)	-3.81 (1.02)	<0.001
Usually sleeps under a net	Individual – KHS	9,763 (57.6)	- 1.78 (0.52)	0.001
Taken anthelmintic in previous 6 months	Individual – KHS	7,494 (44.2)	+ 1.24 (0.48)	0.01
Nutritional proxies				
Wasting in under 5s (%) ²				
Under 5%	Ecological – interpolated	4,027 (23.8)	(vs <5%)	
5-15%	from 2008 DHS	10,257 (60.6)	-5.58 (1.04)	<0.001
Over 15%		2,657 (15.7)	-8.14 (1.21)	<0.001
Low BMI in adult women (%) ³				
Under 20%	Ecological – interpolated	12,867 (76.0)	(vs. <20%)	
Over 20%	from 2008 DHS	4,072 (24.0)	-1.98 (0.80)	0.01
Consumption iron rich foods in under 5s (%) ⁴				
Under 25%	Ecological – interpolated	14,572 (86.0)	(vs. <25%)	
Over 25%	from 2008 DHS	2,369 (14.0)	+ 1.07 (1.24)	0.4
School-feeding (SF) status: ⁵				
Not a SF district	School-level – MoE registers	3,690 (21.8)	(vs. not SF dist)	
SF district – no programme in school		8,417 (49.7)	-7.34 (1.08)	<0.001
SF district – programme in school		4,834 (28.5)	-4.80 (1.19)	<0.001
Generalised livelihood zone ⁶				
Mixed farming		3,295 (19.5)	(vs. mixed)	
Pastoral and agro-pastoral	Ecological – FEWSNET	4,991 (29.5)	+ 1.11 (1.45)	0.4
Riverine and fishing		641 (3.8)	+ 2.08 (2.42)	0.4
Medium potential farming		4,073 (24.0)	- 2.93 (1.13)	0.02
High potential farming		3,941 (23.3)	+ 1.37 (2.14)	0.2
Food security status ⁷				
Generally food secure	Ecological – FEWSNET	2,756 (16.3)	(vs. secure)	
Moderately food insecure		7,755 (45.8)	-3.20 (1.27)	0.01
Highly food insecure		6,048 (35.7)	- 6.17 (1.25)	<0.001
Extremely food insecure		382 (2.3)	- 5.12 (1.52)	0.001

KSH; Kenya School Health survey (conducted 2008-2010). DHS; Kenya Demographic Health Survey conducted 2008/2009. MoE; Kenya Ministry of Education. FEWSNET; Famine early warning system network. SE; standard error.

¹ household head educated to primary complete or above (reported).

^{2,3,4} Interpolated from Kenya 2008/2009 DHS survey

² school included in either the Homegrown School Meals Programme (HGSMP) or Emergency HGSMP 2010 registers

⁶ based on the USAID/FEWSNET Generalised Livelihood Zone 2010 Classification for Kenya

⁷ based on the USAID/FEWSNET Food Insecurity Severity Scale for the quarter corresponding to the KSH survey date.

Formatted: Font: Italic

Table 3 Bayesian hierarchical multivariate regression model for haemoglobin levels.

Variable	Mean Hb shift (g/L)	95% BCI / SD
Baseline haemoglobin	117.3	(113.9, 120.5)
Sex (male)	-1.53	(-1.94, -1.13)
Age (in years)	+1.08	(+0.99, +1.17)
SES quintile (vs. quintiles 2-4)		
lowest	-0.56	(-1.14, +0.02)
highest	0.25	(-0.31, +0.86)
Med/high intensity <i>S. haematobium</i> (>500 epg)	-1.75	(-2.92, -0.60)
Med/high intensity hookworm (>100 epg)	-1.03	(-2.00, -0.02)
<i>Malaria-P. falciparum</i> parasitaemic	-3.00	(-4.29, -1.68)
Med/high hookworm – <i>malaria-P. falciparum</i> co-infection	-5.56	(-8.64, -2.67)
Wasting in under 5s (vs <5% prevalence)		
5-15%	-4.39	(-6.56, -2.09)
>15%	-9.74	(-12.16, -7.45)
School-feeding (SF) status:		
SF district – no programme in school	-2.14	(-4.76, +0.46)
SF district – programme in school	-2.78	(-5.43, -0.21)
Random effect terms		
Individual σ^2	183.6	(2.02)
Spatial σ^2 (school-level)	29.5	(7.93)
Range of spatial correlation (km)	194	(99,318)

Formatted: Font: Italic

Formatted: Font: Italic

Table 4: Predicted adjusted population attributable risk fractions for key predictors across Kenya. Estimates are based on the predicted numbers anaemic in each risk group given by the mean and individual-level standard deviations at each realisation of the Bayesian multivariate model for Hb, and are shown for two demographic groups by Province.

	Population Attributable Risk fractions by Province, % (95% BCI)					
	Central	Coast	Eastern	Nyanza	Rift Valley	Western
14 year old girls:						
Observed prevalence of anaemia	3.3	37.3	34.7	30.8	19.4	25.0
PAF for:						
Med/high intensity hookworm	<i>n/a</i>	13.0 (3.7, 23.1)	5.9 (0, 12.5)	8.2 (0, 30.0)	5.0 (0, 12.5)	29.0 (1.4, 55.1)
Malaria	<i>n/a</i>	5.8 (2.0, 10.4)	<i>n/a</i>	27.6 (0, 49.1)	<i>n/a</i>	23.6 (0, 40.9)
Schistosomiasis	<i>n/a</i>	3.1 (0, 0.08)	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
Acute malnutrition ^a	<i>n/a</i>	18.5 (0, 37.9)	18.4 (0, 73.0)	<i>n/a</i>	16.0 (0, 52.5)	<i>n/a</i>
7 year old boys:						
Observed prevalence of anaemia	5.0	43.3	23.2	43.2	22.4	19.2
PAF for:						
Med/high intensity hookworm	<i>n/a</i>	5.6 (1.0, 10.3)	<i>n/a</i>	11.0 (0, 21.6)	3.4 (0, 6.6)	13.5 (0, 27.5)
Malaria	<i>n/a</i>	1.5 (0.3, 6.1)	<i>n/a</i>	14.3 (0, 32.0)	<i>n/a</i>	15.8 (0, 28.1)
Schistosomiasis	<i>n/a</i>	0 (0, 5.5)	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>	<i>n/a</i>
Acute malnutrition ^a	<i>n/a</i>	16.4 (0, 33.5)	16.4 (0, 67.1)	<i>n/a</i>	14.1 (0, 45.9)	<i>n/a</i>

^a Exposure to "acute malnutrition", prevalence of wasting in under 5s exceeds 5%
n/a – no children are exposed (i.e. prevalence of predictor is zero in this demographic group)