Table 1 Distribution and conditional risk of DSM-IV/CIDI PTSD associated with exposure to the 29 types of traumatic experience (TE) assessed in the WMH Surveys (N=47,566)

	Proportion of all TE exposures % (SE)	Conditional risk of PTSD % (SE)	Proportion of all PTSD % (SE)
Exposed to organized violence			
Civilian in war zone	1.4 (0.1)	1.3 (0.5)	0.5 (0.2)
Civilian in region of terror	1.0 (0.1)	1.6 (0.6)	0.4 (0.1)
Relief worker in war zone	0.3 (0.1)	0.8 (0.7)	0.1 (0.1)
Refugee	0.7 (0.1)	4.5 (2.0)	0.8 (0.4)
Kidnapped	0.4 (0.1)	11.0 (3.0)	1.0 (0.3)
Any	3.9 (0.2)	2.9 (0.5)	2.8 (0.5)
Participated in organized violence			
Combat experience	1.0 (0.1)	3.6 (0.8)	0.9 (0.2)
Witnessed death/serious injury or discovered dead body	16.2 (0.5)	1.3 (0.3)	5.3 (1.0)
Saw atrocities	2.7 (0.3)	5.4 (4.1)	3.7 (2.8)
Accidentally caused death/serious injury	0.7 (0.1)	2.8 (1.0)	0.5 (0.2)
Purposefully caused death/serious injury	0.7 (0.1)	4.0 (3.1)	0.7 (0.5)
Any	21.3 (0.6)	2.1 (0.6)	11.2 (3.1)
Interpersonal violence			
Witnessed physical fights at home as a child	2.4 (0.1)	3.9 (0.7)	2.3 (0.4)
Childhood physical abuse	2.7 (0.1)	5.0 (1.0)	3.4 (0.7)
Beaten by someone else (not spouse/partner)	3.3 (0.2)	2.5 (0.6)	2.1 (0.5)
Mugged or threatened with weapon	8.2 (0.3)	1.8 (0.4)	3.8 (0.8)
Any	16.6 (0.4)	2.8 (0.3)	11.5 (1.3)
Sexual-relationship violence			
Beaten by spouse/partner	1.4 (0.1)	11.7 (1.3)	4.1 (0.5)
Raped	1.8 (0.1)	19.0 (2.2)	8.4 (1.0)
Sexually assaulted	3.2 (0.2)	10.5 (1.5)	8.4 (1.2)
Stalked	2.9 (0.2)	7.6 (2.0)	5.4 (1.4)
Other event	1.4 (0.1)	9.1 (1.0)	3.1 (0.4)
"Private event" (see text)	1.5 (0.1)	9.2 (1.1)	3.5 (0.4)
Any	12.1 (0.3)	10.9 (0.8)	32.9 (2.1)
Other life-threatening TEs			
Life-threatening illness	5.1 (0.2)	2.0 (0.3)	2.5 (0.4)
Life-threatening motor vehicle accident	6.2 (0.2)	2.6 (0.4)	4.1 (0.7)
Other life-threatening accident	3.0 (0.2)	4.9 (2.3)	3.7 (1.8)
Natural disaster	3.9 (0.4)	0.3 (0.1)	0.3 (0.1)
Toxic chemical exposure	3.5 (0.3)	0.1 (0.0)	0.1 (0.0)
Other man-made disaster	1.9 (0.2)	2.9 (1.3)	1.4 (0.7)
Any	23.7 (0.6)	2.0 (0.4)	12.0 (2.1)
Network traumatic experiences			
Unexpected death of loved one	16.8 (0.4)	5.4 (0.5)	22.6 (1.9)
Life-threatening illness of child	3.3 (0.1)	4.8 (0.6)	4.0 (0.5)
Other traumatic experience of loved one	2.4 (0.2)	5.1 (1.3)	3.1 (0.8)
Any	22.5 (0.4)	5.3 (0.4)	29.7 (2.0)
Total	100.0	4.0 (0.2)	100.0

CIDI – Composite International Diagnostic Interview, PTSD – post-traumatic stress disorder, WMH – World Mental Health

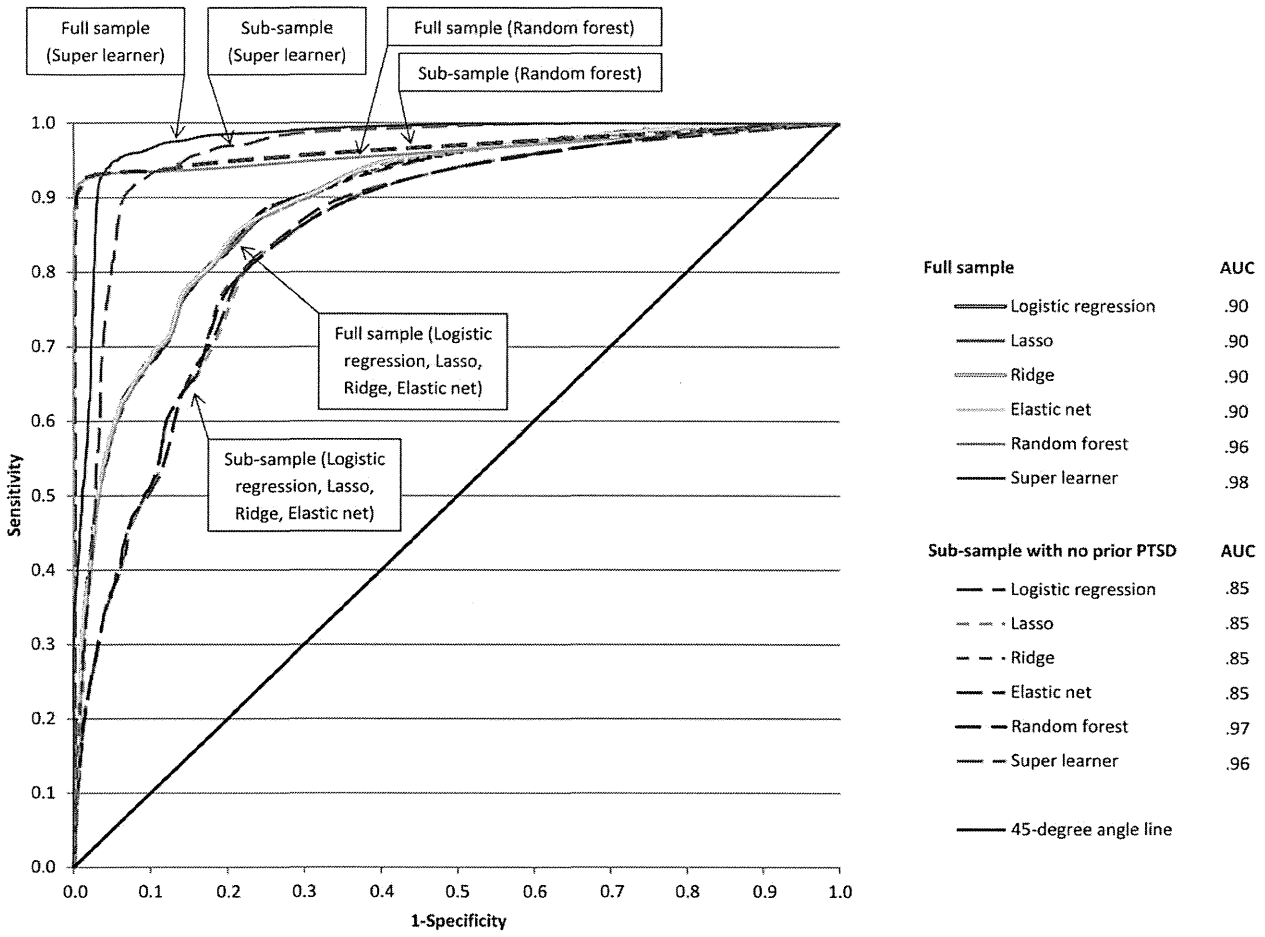

Figure 1 Receiver operating characteristic (ROC) curves for predicted probability of DSM-IV/CIDI PTSD after TE exposure based on the different algorithms in the total sample (N=47,466) and the sub-sample of exposures occurring to respondents with no history of prior PTSD (N=45,556). CIDI – Composite International Diagnostic Interview, TE – traumatic experience, PTSD – post-traumatic stress disorder, AUC – area under the curve

correlated set to zero; the ridge penalty (MPP=0), which uses proportional coefficient shrinkage to retain all predictors; and an intermediate elastic net (MPP=0.5), which combines both approaches. Internal cross-validation was used to select the coefficient in front of the penalty. The algorithms were implemented in the R-package glmnet (24).

Finally, we used an ensembling method known as super learner (28,29) to generate an optimally weighted composite prediction algorithm averaged across the five individual algorithms using internal cross-validation implemented in the R-package Super Learner (30).

It is important to note that the internal cross-validation used in the penalized regressions improves on a more conventional approach, that fits a model in a discovery sample and then tests the model fit in a hold-out sample, in two ways. First, the internal cross-validation used the 10-fold cross-validation technique, which divides the sample into 10 equal-sized sub-samples and estimates a model for each of a large number of fixed coefficients in front of the penalty 10 times, in each of which one of the sub-samples is held out and then the coefficients are applied to the hold-out sample. Model fit was then estimated across the 10 hold-out subsamples to evaluate model fit for the fixed value of the coefficient in front of the penalty. The value of that coefficient was then selected to maximize cross-validated model fit. Second, MPP itself was varied, which leads to variation in the number of predictors in the model. Super learner applied a separate 10-fold cross-validation to this entire set of procedures to assign differential weights to the models with different MPP values as well as to the other algorithms.

Individual-level predicted PTSD probabilities based on the separate algorithms and super learner were created, receiver operating characteristic (ROC) curves generated, and AUC calculated to evaluate prediction accuracy. Super learner predicted probabilities were then discretized into ventiles (20 groups of equal size ordered by percentiles) and cross-classified with observed PTSD. As prior PTSD was a dominant predictor in all algorithms, analysis was replicated for the 45,556 TE exposures that occurred to respondents without a history of prior PTSD. All analyses were based on

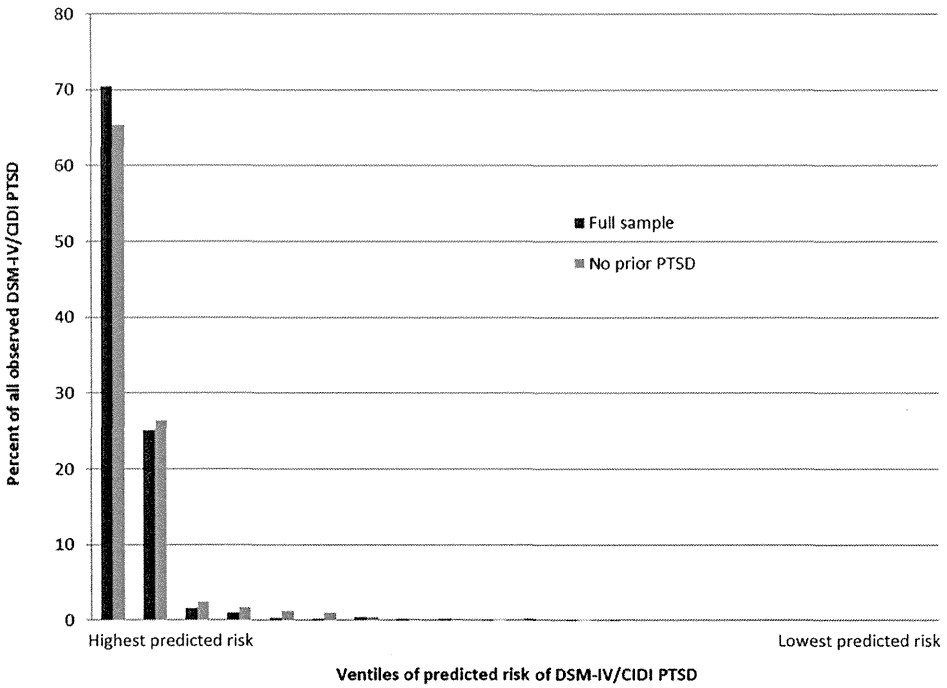

Figure 2 Concentration of risk for DSM-IV/CIDI PTSD. CIDI – Composite International Diagnostic Interview, PTSD – post-traumatic stress disorder

weighted data to adjust for individual differences in probabilities of TE selection.

RESULTS

Distribution and associations of TEs with PTSD

Weighted DSM-IV/CIDI PTSD prevalence was 4.0% in the total sample, and ranged across TEs between 0.1-0.3% (natural and man-made disasters) and 19.0% (rape) (χ^2=639.4, df=28, p<0.001) (Table 1).

The three TEs accounting for the highest proportions of PTSD cases included unexpected death of a loved one, rape, and other sexual assault. Unexpected death of a loved one was the most commonly reported TE (accounting for 16.8% of all TE exposures) and accounted for a somewhat higher proportion of PTSD (22.6%) than of all TE exposures, due to a conditional risk of PTSD slightly higher than average (5.4%). The other two TEs with highest proportions of PTSD cases (8.4% each) were rape and sexual assault. Rape and sexual assault were both much less common than unexpected death of a loved one (rape accounting for 1.8% of all TE exposures and sexual assault for 3.2%), but had much higher conditional PTSD risks (19.0 and 10.5%, respectively).

These high conditional PTSD risks associated with rape and sexual assault were part of a broader pattern of highest PTSD risk being associated with TEs involving interpersonal violence (kidnapping, beaten by spouse/partner, rape, sexual assault), which together accounted for 6.8% of TE exposures and 21.9% of PTSD.

Concentration of risk

ROC curves show that super learner substantially outperformed the individual algorithms other than random forests (AUC=0.96 vs. 0.90 in the total sample; 0.97 vs. 0.85 in the sub-sample with no prior PTSD) (Figure 1).

Inspection of observed PTSD distributions across ventiles of predicted risk based on super learner shows that 95.6% of observed PTSD occurred after the 10% of exposures having highest predicted risk (Figure 2). Conditional PTSD risk was 56.3% in the highest ventile, 20.0% in the second highest ventile, and 0.0-1.3% in the remaining 18 ventiles.

In the sub-sample with no history of prior PTSD, 91.9% of observed PTSD occurred after the 10% of exposures having highest predicted risk. Conditional PTSD risk was 32.2% for the highest ventile, 13.0% for the second highest ventile, and 0.0-1.2% in the remaining 18 ventiles.

Stability of results

Results were found to be stable across sub-samples defined by individual-level characteristics (sex, age, income) and country-level characteristics (economic development, recent history of war or sectarian violence) (Table 2). Between

Table 2 Concentration of observed DSM-IV/CIDI PTSD in the 10% of exposures having highest predicted risk based on the super learner algorithm across sub-samples

	Proportion of all PTSD associated with the 10% of exposures having highest predicted risk		Conditional observed PTSD risk in the 10% of exposures having highest predicted risk vs. other exposures			
			Total sample		Respondents with no prior PTSD	
	Total sample % (SE)	Respondents with no prior PTSD % (SE)	Top 10% % (SE)	Other 90% % (SE)	Top 10% % (SE)	Other 90% % (SE)
Gender						
Male	97.6 (0.6)	96.4 (1.1)	43.4 (3.8)	0.06 (0.01)	27.7 (4.2)	0.05 (0.01)
Female	94.3 (1.0)	92.6 (1.4)	35.9 (1.4)	0.40 (0.07)	21.4 (1.1)	0.34 (0.07)
Age at TE exposure						
Less than 25	95.2 (0.9)	93.3 (1.4)	41.5 (2.4)	0.22 (0.04)	23.0 (2.3)	0.18 (0.03)
25 or older	95.8 (1.2)	94.6 (1.8)	33.9 (2.0)	0.18 (0.05)	23.3 (1.6)	0.16 (0.06)
Education*						
Low/low-average	94.7 (0.8)	92.6 (1.2)	37.5 (1.8)	0.24 (0.04)	21.0 (1.4)	0.20 (0.03)
High-average/high	96.9 (1.4)	96.1 (2.1)	39.5 (3.1)	0.13 (0.06)	28.2 (2.9)	0.11 (0.06)
Country World Bank income level						
High	95.5 (0.6)	94.3 (0.8)	34.3 (1.5)	0.26 (0.03)	20.5 (1.2)	0.20 (0.03)
All others	95.3 (2.0)	92.6 (3.2)	52.6 (3.6)	0.14 (0.06)	33.2 (4.2)	0.13 (0.06)
Country involved in war or sectarian violence**						
Yes	95.9 (1.7)	94.4 (2.5)	44.4 (5.2)	0.13 (0.05)	28.8 (5.2)	0.13 (0.05)
No	95.3 (0.8)	93.6 (1.2)	36.6 (1.3)	0.23 (0.04)	21.6 (1.0)	0.19 (0.04)

PTSD – post-traumatic stress disorder, TE – traumatic experience
*Educational level relative to others in the same country; **countries classified "yes" include Colombia, Israel, Lebanon, Nigeria, Northern Ireland, and South Africa

94.3% and 97.6% of observed PTSD in each sub-sample was associated with the 10% of TEs having highest predicted risk (92.6-96.4% in the sub-sample with no prior PTSD). PTSD prevalence in these high-risk sub-samples was 33.9-52.6% (20.5-33.2% in the sub-sample with no prior PTSD).

Components of risk

Although it is hazardous to interpret individual machine learning model coefficients, since the algorithms maximize overall prediction accuracy at the expense of individual coefficient accuracy, an understanding of important predictors is nonetheless useful. We used a two-part approach to achieve this understanding. We first examined multivariate predictor profiles in 100 bootstrapped r-part trees to determine which interactions were important. These profiles all involved history of prior PTSD interacting either with history of prior unexpected death of a loved one and/or with history of prior sexual trauma. We then included dummy variables for those multivariate profiles along with variables for the main effects of individual predictors in lasso regressions.

Socio-demographic differences in PTSD risk were restricted to elevated odds ratios among women (1.5-1.6) and the previously married (1.5). Nine TE types also had elevated odds ratios, all but one of which (the exception being exposure to a life-threatening accident, with odds ratios of 1.4-1.8) involved interpersonal violence: rape (3.2-3.5), kidnap (1.8-3.4), childhood physical abuse (1.5-1.8), witnessing atrocities (1.4), and four other (than rape) TE types in the relationship-sexual violence factor (1.5-1.8). Three TE types had meaningfully reduced odds ratios: witnessing death/injury, toxic chemical spill, and natural disasters (0.4-0.7) (Table 3).

Five summary measures of collateral TEs occurring in the same year as the focal TE had meaningful odds ratios. Four of these five (the exception being unexpected death of a loved one) involved violence: two or more TEs in the organized violence factor, three or more TEs in the interpersonal violence factor, and any as well as two or more TEs in

Table 3 Lasso penalized logistic regression coefficients (odds ratios) to predict onset of DSM-IV/CIDI PTSD after exposure to a traumatic experience (TE)

	Total sample OR	Sub-sample without prior PTSD OR	Total sample with interaction OR
Focal TE			
Organized violence			
Witnessed atrocities	1.4	1.4	1.4
Witnessed death/injury or discovered dead bodies	0.6	0.7	0.6
Kidnapped	3.0	1.8	3.4
Interpersonal violence			
Childhood physical abuse	1.5		
Sexual-relationship violence			
Beaten by spouse-partner	1.8	1.5	1.7
Raped	3.2	3.5	3.5
Sexually assaulted			1.5
"Private TE" (see text)	1.8	1.5	1.8
Some other TE	1.6		1.5
Other			
Toxic chemical exposure	0.5	0.6	0.5
Natural disaster	0.4		0.5
Other life-threatening accident	1.8	1.4	1.6
Collateral TEs			
Multiple (2+) participants in organized violence	2.5	3.8	2.7
High (3+) exposure to interpersonal violence	6.8	11.5	9.3
Any exposure to sexual-relationship violence	1.6	1.7	1.6
Multiple (2+) exposures to sexual-relationship violence	4.0	3.2	3.5
Unexpected death of loved one	2.1	2.1	2.2
Socio-demographics			
Female		1.6	1.5
Previously married	1.5		1.5
Lifetime prior TEs			
Witnessed atrocities	1.6	1.8	1.6
Raped		2.3	1.6
Sexually assaulted			1.4
Unexpected death of loved one			1.6
Lifetime prior DSM-IV/CIDI disorders			
Separation anxiety disorder	2.0	1.7	1.9
Specific phobia			1.5
Attention-deficit/hyperactivity disorder	1.6		1.7
Generalized anxiety disorder	2.2	2.5	2.2
PTSD	27.2		
Interactions of lifetime prior PTSD with lifetime prior			
Sexual violence and unexpected death of loved one			5.4
Sexual violence but no unexpected death of loved one			12.9
Unexpected death of loved one but no sexual violence			5.7
Neither sexual violence nor unexpected death of loved one			134.7

CIDI – Composite International Diagnostic Interview, PTSD – post-traumatic stress disorder, OR – odds ratio

the relationship-sexual violence factor. The collateral TEs involving single exposures (sexual or death) had odds ratios of 1.6-2.2, while those involving two-three or more exposures had odds ratios of 2.5-11.5. Four of the 29 prior lifetime TE types had meaningful odds ratios, all of them elevated: witnessing atrocities (1.6-1.8), being raped (1.6-2.3) or sexually assaulted (1.4), and experiencing the unexpected death of a loved one (1.6) (Table 3).

Five of the 14 prior lifetime DSM-IV/CIDI disorders had meaningful odds ratios: ADHD (1.6-1.7), separation anxiety disorder (1.7-2.0), specific phobia (1.5), generalized anxiety disorder (2.2-2.5), and PTSD (27.2) (Table 3). The high odds ratio for prior PTSD was due to the 3.5% of exposures occurring to respondents with prior PTSD accounting for 40.5% of all episodes of PTSD. Disaggregation into underlying multivariate profiles showed that this strong effect of prior PTSD was limited to people marked as vulnerable by virtue of having past PTSD associated with TEs generally not associated with high PTSD risk.

DISCUSSION

Although caution is needed in interpreting these results, since the WMH Survey data were based on retrospective reports and fully structured lay-administered diagnostic interviews, it is nonetheless striking that we were able to produce a prediction algorithm in which the vast majority of PTSD cases were associated with the 10% of TE exposures having highest predicted risk. By far the most powerful predictor in the algorithm was history of prior PTSD, but a number of other prior lifetime mental disorders were also significant predictors, along with a number of measures of prior lifetime trauma exposure as well as socio-demographic characteristics of respondents and information about characteristics of the focal TE.

Limitations introduced by the retrospective nature of the data could have led to upward bias in odds ratios if respondents defined as having a history of PTSD were more likely than others to recall prior lifetime TE exposures and/or mental disorders. Importantly, evidence has been presented in the literature that this type of bias does, in fact, exist in retrospective reports among people with PTSD (31-33). In addition, the concentration-of-risk estimates could have been upwardly biased compared to those that would be found among people who sought help in the immediate aftermath of TE exposure in criminal justice or health care settings, to the extent that this help-seeking predicted subsequent PTSD.

Despite these limitations, the WMH Survey results are important in suggesting that a PTSD risk algorithm based on machine learning methods might help improve targeting and subsequent cost-effectiveness of preventive interventions for PTSD by pinpointing the small proportion of TE-exposed people having high PTSD risk that account for most subsequent PTSD. Our study is much larger than all other previous studies attempting to predict onset of PTSD from information about trauma types and pre-trauma predictors. We showed that a composite risk score can be constructed from such data that classifies the vast majority of people who go on to develop PTSD into a high-risk segment of the population.

External validation of the risk algorithms in prospective samples is, however, needed. A number of such prospective studies exist. All these studies are much smaller than the WMH Survey database and not all assessed the full range of predictors considered in our analysis. Nonetheless, the strong results found here suggest that it would be valuable to carry out replications in these prospective studies. We are currently involved in several collaborative replication analyses of this sort and are eager to work with others to evaluate the extent to which our algorithm fits in independent prospective samples. In addition, we are interested in collaborating with other groups to apply the methods used here to determine the predictive accuracy of algorithms based on data involving an expanded set of predictors, including biomarkers. If the results of this ongoing work are encouraging, subsequent prospective studies should be designed so that they include the full range of predictors found to be important. This would advance the agenda of creating broadly useful PTSD risk algorithms (and subsequent algorithms to predict other psychopathological responses to TE exposure) to target preventive interventions across a wide range of settings.

It is important to note that different predictors will almost certainly be found to be important in different populations (e.g., military personnel, first responders in disaster situations, civilians in less developed countries), in association with different TEs (e.g., sectarian violence in war zones or regions of terror, large-scale natural disasters) and in different screening settings (e.g., temporary emergency clinics established at the site of natural or man-made disasters, medical clinics in war zones, trauma units, emergency departments) (6,7,11,13). This means that an expansion of the current line of work will ultimately lead either to the development of a family of risk prediction algorithms or to a consolidated master algorithm that allows for complex interactions across populations, TEs, and screening settings.

It will also be important, in developing such algorithms, to be sensitive to variation in the costs of collecting different types of data (e.g., self-reports versus biomarkers), even data ostensibly assessing the same underlying constructs (e.g., self-reports of impulsivity versus neurocognitive tests of impulsivity), as well as the burdens associated with administering detailed risk factors assessments (both in terms of time burden and the psychological burden of asking people detailed questions of a sensitive nature in the immediate aftermath of TE exposure). Thoughtful analysis will be needed of the cost-benefit trade-offs associated with including-excluding expensive and burdensome elements of the assessments depending on strength of predictions.

Acknowledgements

The research reported here is supported by U.S. Public Health Service grant 1 R01MH101227-01A1 and is carried out in conjunction with the WHO's WMH Survey initiative, which is supported by the National Institute of Mental Health (R01 MH070884 and R01 MH093612-01), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation, the U.S. Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558), the Fogarty International Center (FIRCA R03-TW006481), the Pan American Health Organization, Eli Lilly and Company, Ortho-McNeil Pharmaceutical, GlaxoSmithKline, and Bristol-Myers Squibb. The authors thank the staff of the WMH Data Collection and Data Analysis Coordination Centres for assistance with instrumentation, fieldwork, and consultation on data analysis. None of the funders had any role in the design, analysis, interpretation of results, or preparation of this paper. A complete list of all within-country and cross-national WMH Survey publications can be found at http://www.hcp.med.harvard.edu/WMH/. The views and opinions expressed in this report are those of the authors and should not be construed to represent the views of the sponsoring organizations, agencies, or governments.

References

1. Kessler R. Posttraumatic stress disorder: the burden to the individual and to society. J Clin Psychiatry 2000;61(Suppl. 5):4-12.
2. Forneris CA, Gartlehner G, Brownley KA et al. Interventions to prevent post-traumatic stress disorder: a systematic review. Am J Prev Med 2013;44:635-50.
3. Kliem S, Kroger C. Prevention of chronic PTSD with early cognitive behavioral therapy. A meta-analysis using mixed-effects modeling. Behav Res Ther 2013;51:753-61.
4. Shalev AY, Ankri Y, Israeli-Shalev Y et al. Prevention of posttraumatic stress disorder by early treatment: results from the Jerusalem Trauma Outreach And Prevention study. Arch Gen Psychiatry 2012; 69:166-76.
5. Ostrowski SA, Delahanty DL. Prospects for the pharmacological prevention of post-traumatic stress in vulnerable individuals. CNS Drugs 2014;28:195-203.
6. Brewin CR, Andrews B, Valentine JD. Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. J Consult Clin Psychol 2000;68:748-66.
7. Ozer EJ, Best SR, Lipsey TL et al. Predictors of posttraumatic stress disorder and symptoms in adults: a meta-analysis. Psychol Bull 2003;129:52-73.
8. Tolin DF, Foa EB. Sex differences in trauma and posttraumatic stress disorder: a quantitative review of 25 years of research. Psychol Bull 2006;132:959-92.
9. Breslau N. Epidemiologic studies of trauma, posttraumatic stress disorder, and other psychiatric disorders. Can J Psychiatry 2002; 47:923-9.
10. DiGangi JA, Gomez D, Mendoza L et al. Pretrauma risk factors for posttraumatic stress disorder: a systematic review of the literature. Clin Psychol Rev 2013;33:728-44.
11. Heron-Delaney M, Kenardy J, Charlton E et al. A systematic review of predictors of posttraumatic stress disorder (PTSD) for adult road traffic crash survivors. Injury 2013;44:1413-22.
12. Roberts AL, Gilman SE, Breslau J et al. Race/ethnic differences in exposure to traumatic events, development of post-traumatic stress disorder, and treatment-seeking for post-traumatic stress disorder in the United States. Psychol Med 2011;41:71-83.
13. van Zuiden M, Kavelaars A, Geuze E et al. Predicting PTSD: pre-existing vulnerabilities in glucocorticoid-signaling and implications for preventive interventions. Brain Behav Immun 2013;30:12-21.
14. Baca-Garcia E, Perez-Rodriguez MM, Basurte-Villamor I et al. Using data mining to explore complex clinical decisions: a study of hospitalization after a suicide attempt. J Clin Psychiatry 2006; 67:1124-32.
15. Connor JP, Symons M, Feeney GF et al. The application of machine learning techniques as an adjunct to clinical decision making in alcohol dependence treatment. Subst Use Misuse 2007;42:2193-206.
16. World Bank. Data: countries and economies. http://data.worldbank.org/country.
17. Heeringa SG, Wells EJ, Hubbard F et al. Sample designs and sampling procedures. In: Kessler RC, Üstün TB (eds). The WHO World Mental Health Surveys: global perspectives on the epidemiology of mental disorders. New York: Cambridge University Press, 2008:14-32.
18. Kessler RC, Üstün TB. The World Mental Health (WMH) Survey Initiative Version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). Int J Methods Psychiatr Res 2004;13:93-121.
19. Haro JM, Arbabzadeh-Bouchez S, Brugha TS et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. Int J Methods Psychiatr Res 2006; 15:167-80.
20. First MB, Spitzer RL, Gibbon M et al. Structured Clinical Interview for DSM-IV Axis I Disorders, Research Version, Non-patient Edition (SCID-I/NP). New York: Biometrics Research, New York State Psychiatric Institute, 2002.
21. Knäuper B, Cannell CF, Schwarz N et al. Improving accuracy of major depression age-of-onset reports in the US National Comorbidity Survey. Int J Methods Psychiatr Res 1999;8:39-48.
22. Hosmer DW, Lemeshow S. Applied logistic regression, 2nd ed. New York: Wiley, 2001.
23. Breiman L. Random forests. Mach Learn 2001;45:5-32.
24. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw 2010; 33:1-22.
25. Berk RA. Statistical learning from a regression perspective. New York: Springer, 2008.
26. Liaw A, Wiener M. Classification and regression by randomForest. R News 2002;2:18-22.
27. Thernau T, Atkinson B, Ripley B. Rpart: recursive partitioning. R Package 4.1-0., 2012.
28. van der Laan MJ, Polley EC, Hubbard AE. Super learner. Stat Appl Genet Mol Biol 2007;6:Article 25.
29. van der Laan MJ, Rose S. Targeted learning: causal inference for observational and experimental data. New York: Springer, 2011.
30. Polley EC, van der Laan MJ. SuperLearner: Super Learner prediction, Package Version 2.0-4. Vienna: R Foundation for Statistical Computing, 2011.
31. Harvey AG, Bryant RA. Memory for acute stress disorder symptoms: a two-year prospective study. J Nerv Ment Dis 2000;188: 602-7.
32. Roemer L, Litz BT, Orsillo SM et al. Increases in retrospective accounts of war-zone exposure over time: the role of PTSD symptom severity. J Trauma Stress 1998;11:597-605.
33. Zoellner LA, Foa EB, Brigidi BD et al. Are trauma victims susceptible to "false memories"? J Abnorm Psychol 2000;109:517-24.

DOI 10.1002/wps.20150

Published in final edited form as:
Lancet. 2014 September 13; 384(9947): 1005–1070. doi:10.1016/S0140-6736(14)60844-8.

Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013

A full list of authors and affiliations appears at the end of the article.

Summary

Background—The Millennium Declaration in 2000 brought special global attention to HIV, tuberculosis, and malaria through the formulation of Millennium Development Goal (MDG) 6. The Global Burden of Disease 2013 study provides a consistent and comprehensive approach to disease estimation for between 1990 and 2013, and an opportunity to assess whether accelerated progress has occurred since the Millennium Declaration.

Methods—To estimate incidence and mortality for HIV, we used the UNAIDS Spectrum model appropriately modified based on a systematic review of available studies of mortality with and without antiretroviral therapy (ART). For concentrated epidemics, we calibrated Spectrum models to fit vital registration data corrected for misclassification of HIV deaths. In generalised epidemics, we minimised a loss function to select epidemic curves most consistent with prevalence data and demographic data for all-cause mortality. We analysed counterfactual scenarios for HIV to assess years of life saved through prevention of mother-to-child transmission (PMTCT) and ART. For tuberculosis, we analysed vital registration and verbal autopsy data to estimate mortality using cause of death ensemble modelling. We analysed data for corrected case-notifications, expert opinions on the case-detection rate, prevalence surveys, and estimated cause-specific mortality using Bayesian meta-regression to generate consistent trends in all parameters.

Correspondence to: Christopher J L Murray.

Contributors: CJLM, ADL, and TV prepared the first draft. CJLM, KFO, CG, SSL, TMW, DAR, EAD, NG, RMB, JCB, HCD, LD, JAS, DEP, TDF, BKP, EKJ, MSC, ADL, and TV finalised the draft based on comments from other authors and reviewer feedback. CJLM, SSL, and TV had the idea for the study and provided overall guidance. CJLM, KFO, CG, SSL, TMW, DAR, EAD, NG, HW, MN, DD, KRH, KF, DEP, TDF, ADF, and TV did all modeling. CJLM, KFO, CG, SSL, TMW, DAR, EAD, NG, RMB, HW, HCD, DD, DEP, ADF, and TV did the statistical analysis of model results. All other authors provided data, developed models, reviewed results, initiated modeling infrastructure, and reviewed the paper.

†Authors listed alphabetically

‡Joint senior authors

Declaration of interests: CC has received consultancy, lecture fees and honoraria from AMGEN, GSK, Alliance for Better Bone Health, MSD, Eli Lilly, Pfizer, Novartis, Servier, Merck, Medtronic and Roche. LD has been funded by and acted as a consultant to UNAIDS, UNODC, and WHO. BD Gessner works for AMP which receives grant specific support from Crucell, GSK, Novartis, Pfizer, and SanofiPasteur and unrestricted support from SanofiPasteur all for work on vaccines and vaccine preventable diseases. Related to the current work, AMP receives grant support from GSK for work on issues related to malaria vaccines. KBG received the NHMRC-Gustav Nossal scholarship sponsored by CSL in 2012. This award is peer-reviewed through the standard NHMRC peer-review process; CSL played no part in selection of the awardee. PJ is supported by a career development fellowship from the Wellcome Trust, Public Health Foundation of India and a Consortium of United Kingdom Universities. Walter Mendoza is program analyst at the UNFPA country office in Peru, which not necessarily endorses the study. JAS has received research grants from Takeda and Savient and consultant fees from Savient, Takeda, Regeneron and Allergan. JAS is a member of the executive of OMERACT, an organization that develops outcome measures in rheumatology and receives arms-length funding from 36 companies; a member of the American College of Rheumatology's Guidelines Subcommittee of the Quality of Care Committee; and a member of the Veterans Affairs Rheumatology Field Advisory Committee.

Europe PMC Funders Author Manuscripts

We analysed malaria mortality and incidence using an updated cause of death database, a systematic analysis of verbal autopsy validation studies for malaria, and recent studies (2010–13) of incidence, drug resistance, and coverage of insecticide-treated bednets.

Findings—Globally in 2013, there were 1·8 million new HIV infections (95% uncertainty interval 1·7 million to 2·1 million), 29·2 million prevalent HIV cases (28·1 to 31·7), and 1·3 million HIV deaths (1·3 to 1·5). At the peak of the epidemic in 2005, HIV caused 1·7 million deaths (1·6 million to 1·9 million). Concentrated epidemics in Latin America and eastern Europe are substantially smaller than previously estimated. Through interventions including PMTCT and ART, 19·1 million life-years (16·6 million to 21·5 million) have been saved, 70·3% (65·4 to 76·1) in developing countries. From 2000 to 2011, the ratio of development assistance for health for HIV to years of life saved through intervention was US$4498 in developing countries. Including in HIV-positive individuals, all-form tuberculosis incidence was 7·5 million (7·4 million to 7·7 million), prevalence was 11·9 million (11·6 million to 12·2 million), and number of deaths was 1·4 million (1·3 million to 1·5 million) in 2013. In the same year and in only individuals who were HIV-negative, all-form tuberculosis incidence was 7·1 million (6·9 million to 7·3 million), prevalence was 11·2 million (10·8 million to 11·6 million), and number of deaths was 1·3 million (1·2 million to 1·4 million). Annualised rates of change (ARC) for incidence, prevalence, and death became negative after 2000. Tuberculosis in HIV-negative individuals disproportionately occurs in men and boys (versus women and girls); 64·0% of cases (63·6 to 64·3) and 64·7% of deaths (60·8 to 70·3). Globally, malaria cases and deaths grew rapidly from 1990 reaching a peak of 232 million cases (143 million to 387 million) in 2003 and 1·2 million deaths (1·1 million to 1·4 million) in 2004. Since 2004, child deaths from malaria in sub-Saharan Africa have decreased by 31·5% (15·7 to 44·1). Outside of Africa, malaria mortality has been steadily decreasing since 1990.

Interpretation—Our estimates of the number of people living with HIV are 18·7% smaller than UNAIDS's estimates in 2012. The number of people living with malaria is larger than estimated by WHO. The number of people living with HIV, tuberculosis, or malaria have all decreased since 2000. At the global level, upward trends for malaria and HIV deaths have been reversed and declines in tuberculosis deaths have accelerated. 101 countries (74 of which are developing) still have increasing HIV incidence. Substantial progress since the Millennium Declaration is an encouraging sign of the effect of global action.

Funding—Bill & Melinda Gates Foundation.

Introduction

The Millennium Declaration in 2000 brought special global attention to HIV, tuberculosis, and malaria through the formulation of Millennium Development Goal 6 (MDG 6). The high priority status of these three diseases in the development community was confirmed through the creation of the Global Fund to Fight AIDS, Tuberculosis and Malaria in 2002. Bilateral initiatives such as the President's Emergency Plan for AIDS Relief and the President's Malaria Initiative also added substantial new resources. From 2000 to 2011, multilaterals, bilaterals, foundations, and non-governmental organisations have invested US $51·6 billion for HIV, $11·3 billion for malaria, and $8·3 billion for tuberculosis (price in 2011 US dollars) in development assistance for health (DAH).[1] Substantial benefits of these

investments have been documented in several studies.[2-9] In the lead up to the MDG deadline of 2015 and amid the global debate on development goals post-2015, important questions have been raised about the advantages and disadvantages of maintaining focus on these three diseases.[5,10-18] The rise of the importance of non-communicable diseases in some regions of the developing world[19-26] have led to calls for goals that cover a broader range of diseases.[10,15,17,19,20,24,25,27,28] At the same time, ambitious goals of zero tuberculosis incidence and deaths and zero HIV incidence and deaths have been formulated by some groups;[29-34] the Secretary-General of the UN had already established a goal of zero malaria deaths by 2015.[35] Understanding the distribution and trends of these three diseases and how they have been affected by the MDG era is an important input to this wider debate.[12,36]

Because of their prominence, there are major UN efforts on an annual basis to track the epidemiology of these three diseases. UNAIDS now produces country estimates of HIV incidence, prevalence, and death every year.[37] Over many years, they have developed a sophisticated modelling approach to track the epidemic—their primary input in generalised epidemics is annual antenatal clinic sero-surveillance data and periodic household surveys that include blood testing.[38,39] The annual Global Tuberculosis Report from the World Health Organization (WHO) provides estimates of incidence and deaths from tuberculosis by country. Crucial inputs to the assessment of incidence are case notifications and national expert opinion on the case-detection rate, and separate modelling of cause of death data from vital registration systems and verbal autopsy studies. For the World Malaria Report, WHO uses a complicated strategy to estimate incidence and mortality that varies by region and age group. For child malaria deaths in sub-Saharan Africa, the main inputs are verbal autopsy studies and estimated malaria risk. Estimates are adjusted post-hoc for coverage of insecticide-treated bednets (ITNs). Outside of sub-Saharan Africa and for low-transmission countries in Africa reported case numbers are combined with an assumed case-fatality rate. These three efforts have provided important insights into the geographical distribution and likely trends in the diseases.

Despite these efforts, extraordinary uncertainty exists at the country-level in the burden of all three. The burdens of HIV and malaria are concentrated in sub-Saharan Africa; countries that, other than South Africa, have very poor vital registration and incomplete notification systems. Tuberculosis is concentrated in Asia and southern Africa where a few more countries have better data systems but there are still huge gaps in information. Modelling strategies for tracking the diseases have evolved to be necessarily complex in view of the incomplete and often conflicting nature of the data. For HIV and malaria, UN modelling efforts explicitly use information about intervention delivery and assumed benefits of intervention. The distinction between data for disease outcomes and data for intervention coverage driving the results of these efforts is blurred. In the more complex modelling strategies, the compounded effect of uncertainty about different parameters can be hard to characterise. Efforts to model the three diseases are largely independent of each other—the exception is recent coordinated efforts to understand the intersection of tuberculosis and HIV.[40]

The Global Burden of Disease 2010 Study provided a comprehensive update on levels and trends of a large number of diseases, injuries, and risk factors for 187 countries from 1990 to

2010.[41-48] The Global Burden of Disease collaboration is now generating annual updates, the first of which is the Global Burden of Disease, Injuries, and Risk Factors Study 2013 (GBD 2013). The GBD 2013 provides an opportunity to examine the evidence on the levels and trends in the three MDG 6 diseases within the comprehensive and coherent framework of the GBD. Compared with GBD 2010, we have given special emphasis in the GBD 2013 to incorporate new data, to more rigorously identify and incorporate further key sources of uncertainty, and to incorporate adjustments for the biases that are present in different data sources. A crucial aspect of the GBD effort is to quantify time trends; comparing trends from 1990 to 2000 and from 2000 to 2013 provides an opportunity to see if there has been accelerated progress since the Millennium Declaration. The GBD 2013 supersedes all previously published GBD results.

Methods

Overview

The overall conceptual and analytical framework for the GBD is described elsewhere.[41-48] Major refinements of the analytical approach for different diseases and risk factors are explored in other papers.[49-51] We summarise here the methods used for the analysis of the three diseases, emphasising refinements since the GBD 2010. All refinements in methods have been applied to the full 1990–2013 time series to ensure comparability of results. Metadata for input sources used in the GBD 2013 will be available in the Global Health Data Exchange on publication of the full GBD 2013 results.

HIV

For the GBD 2010, we primarily used estimates of prevalence and mortality developed by UNAIDS. The main modification was the requirement that the sum of cause-specific mortality from each cause in a country, age, sex, and year group equalled the estimate of all-cause mortality for that country, age, and sex group generated through the analysis of demographic sources.[43] Modifications of HIV deaths through this internal consistency process did not lead to revisions of incidence or prevalence for HIV in a particular age-sex-country-year. For the GBD 2013, we have sought to develop a set of estimates of incidence, prevalence, and mortality from HIV that are internally consistent with each other and also meet the GBD requirement that the sum of each cause of death equals all-cause mortality. Internally consistent means the incidence, prevalence, and death figures are mathematically possible given that prevalence is a function of past incidence, remission, and death rates for any age cohort.

Modified Spectrum Model—UNAIDS uses two key analytical components in their epidemiological estimation. The Estimation and Projections Package (EPP) is used to estimate incidence trajectories that are consistent with prevalence surveys and other prevalence measurements such as antenatal clinic serosurveillance.[52] Spectrum is a compartmental HIV progression model used to generate age-specific incidence, prevalence, and death rates from the EPP incidence curves and assumptions about intervention scale-up and local variation in epidemiology.[53] We have recoded Spectrum in Python, an open source higher level language that can easily be run on a parallelised computational cluster,

following the exact structure of Spectrum to facilitate faster computation required for the uncertainty analysis and consistency analysis with all-cause mortality. We have also made four important modifications to the assumptions.

First, we have altered the Spectrum assumptions about mortality without antiretroviral therapy (ART). Following UNAIDS assumptions, mortality is modelled as shown in figure 1. The death and progression rates between CD4 categories vary by age according to four age-groups, 15–24 years, 25–34 years, 35–44 years, and 45 years or older. UNAIDS estimates a single set of progression and death rates by first fitting a Weibull distribution to data from three east African seroconverter cohorts from the ALPHA network and one miners cohort from South Africa,[54] and then selecting a set of progression and death rates in Excel Solver that minimises the sum of the squared differences between the predicted and Weibull survival probabilities.[55,56] Uncertainty in their estimates for this component was approximated by assuming a coefficient of variation of 0·05 for each mortality rate. To better characterise uncertainty in the progression and death rates, we systematically reviewed the literature on mortality without ART. We searched terms related to pre-ART or ART-naive survival since seroconversion—exact search terms are in the appendix (p 132). After screening, we identified 13 cohort studies that included the cohorts used by UNAIDS from which we extracted survival at each 1-year point after infection. We modelled the logit of the conditional probability of death between years in these studies using the following formula:

$$logit\,(m_{ijk}) = \beta_0 + \sum_{i=1}^{4} \beta_{1i} a_i + \sum_{j=1}^{12} \beta_{2j} t_j + u_k + \varepsilon_{ijk}$$

In the formula, m is conditional probability of death from year t_j to t_{j+1}, a_i is an indicator variable for age group at seroconversion (15–24 years, 25–34 years, 35–44 years, and 45 years or older), t_j is an indicator variable of year since seroconversion, and u_k is a study-level random effect. By sampling the variance-covariance matrix of the regression coefficients and the study level random effect, we generated 1000 survival curves for each age group that capture the systematic variation in survival across the available studies (figure 2, appendix p 18). Across all age groups, median survival ranges from 3·6 years to 29·5 years. For each of the 1000 survival curves, we use the UNAIDS optimisation framework to find a set of progression and death rates that minimises the sum of the squared errors for the fit to the survival curve. For the death rates generated from the optimisation, the coefficient of variation across the set of 1000 is 0·44–1·00 depending on the age-group, which is substantially higher than the UNAIDS assumption of 0·05.

Second, for mortality on ART, UNAIDS used data from five regional cohorts from the IeDEA network to directly estimate death rates by age, sex, and CD4 count, which have been used as the default for all countries in a region.[57,58] Through their country consultation process some of these defaults have been modified. For example, Myanmar assumes a constant mortality rate by initial CD4 group, without any variation by age, sex, or time on treatment. To better characterise real variation in the death rates on ART across programmes, we searched the published literature. Using the terms "HIV", "mortality", and

"antiretroviral therapy" we identified 4996 titles. Screening the abstracts and papers yielded 102 total papers for extraction (appendix p 133). These included mortality and loss to follow-up data for 80 cohorts, age hazard ratios for 40 cohorts, and sex hazard ratios for 86 cohorts. We corrected reported probabilities of death for loss to follow-up using an update of the approach developed by Verguet and colleagues.[59] Verguet and colleagues used tracing and follow-up studies to empirically estimate the relation between death in loss to follow-up and the rate of loss to follow-up. We used DisMod-MR 2.0 to do a meta-regression of the data for on-ART mortality by initial CD4 count separately for high-income countries, GBD developing countries outside of sub-Saharan Africa, and sub-Saharan Africa (appendix p 7). We meta-analysed region-specific age hazard ratios using DisMod-MR 2.0, and region-specific sex hazard ratios using the Stata command *metan*. The age and sex hazard ratios were applied to the CD4-specific mortality rates, accounting for the distribution of ages and sexes in the mortality data. We used 1000 draws from the posterior distribution for each age, sex, and CD4 category for conditional probabilities of death for 0–6 months, 7–12 months, and 13–24 months after initiation of ART as inputs to Spectrum. Table 1 shows HIV-specific mortality rates for people aged 25–34 years on ART in sub-Saharan Africa (see appendix pp 136–38 for HIV-specific mortality for other age groups and regions).

Third, to better capture variation in the age-pattern of incidence, we used the UNAIDS distributions of the relative incidence by age prepared for UNAIDS based on selected cohort studies. To capture the possibility that there is greater variation across countries in the age incidence pattern than in these studies, we increased the uncertainty ranges by an arbitrary 50%.

Fourth, for all other input parameters including the number of individuals receiving ART, prevention of mother-to-child transmission (PMTCT), or co-trimoxazole prophylaxis, we randomly sampled a uniform distribution from 0·9 to 1·1 and used the draw to adjust each parameter. For the sex ratio of incidence, we sampled a wider but arbitrary range from 0·8 to 1·2, because the demographic data in many generalised epidemics indicate that there is less of a difference between the sexes than seen in the population prevalence surveys.

Generalised epidemics—UNAIDS identified 41 countries as generalised epidemics; this distinction is important because for these epidemics the primary sources of information about prevalence come from antenatal clinic serosurveillance and household surveys. In addition to these 41 countries, we have included in this category Senegal, Niger, and India because of the availability of population-based surveys. Prevalence data from countries with generalised epidemics has been analysed by UNAIDS using EPP to generate 1000 samples of incidence curves for people aged 15–49 years consistent with the prevalence data. For each of these 1000 incidence curves, we randomly sampled the parameter distributions for all input parameters ten times to generate 10 000 epidemic curves of incidence, prevalence, and death by age and sex. The selection of 10 000 iterations was based on testing that it would ensure stable uncertainty intervals. Some of these 10 000 death curves exceed in one or more age-sex-year groups the estimate of all-cause mortality based on demographic sources. Because the demographic estimates of all-cause mortality are based on substantial empirical data, these HIV epidemic curves are unlikely to represent reality. These mismatches occur more often in southern Africa. We identified the 250 modified Spectrum

curves and all-cause mortality curves that are most consistent with each other. We define a loss function using the following formula:

$$e_r = \sum_t \sum_a \sum_s max\left(0, m^{HIV}_{r,t,a,s} - 0 \cdot 8^* m^{all-cause}_{r,t,a,s}\right)$$

For run r of a given country, excess mortality, e, is equal to the sum of all non-zero differences between HIV mortality, m^{HIV}, and 0·8 times a randomly selected all-cause mortality draw, $m^{all\text{-}cause}$, across all year-age-sex combinations (t, a, and s, respectively). We compared the Spectrum estimates to 0·8 times the all-cause estimates because this is the highest recorded fraction of deaths in any age-group in any country's cause of death data due to HIV. We randomly paired each of the 10 000 modified spectrum outputs with one of the 1000 all-cause mortality curves generated from the demographic analysis. The 250 pairs that minimised the loss function were selected. When more than 250 have a loss function equal to zero, we randomly chose among this set. We resampled the 250 draws to 1000; 250 are used with resampling for computational convenience. The appendix (pp 18–20) shows mortality at ages 15–59 years from the full set of 10 000 modified spectrum models and the subset that is selected through the matching process for Uganda, South Africa, and Ghana. Demographic data matching selects in South Africa epidemic curves that are at the low end of the distribution with longer median survival; in Uganda this effect is slightly less pronounced and in Ghana pre-matching and post-matching were identical.

Concentrated epidemics with vital registration data—UNAIDS estimates for concentrated epidemics depend critically on two inputs: first, the assessment of prevalence of HIV in high-risk groups (people who inject drugs, men who have sex with men, and female sex workers) for which there is much information in many countries;[60-72] and second, assumptions on the percentage of the population in high-risk groups. Although there is guidance on measurement,[73] real data in most countries are limited. In many countries, UNAIDS estimates are based on local opinion. Resulting assumptions have been highly variable across countries. For example, Uruguay defines 4·5% of its population as men who have sex with men while neighbouring Argentina defines only 1% of its population as men who have sex with men. An alternative source in many countries is cause of death data from national vital registration systems. To track the epidemic using cause of death data can require up to three important adjustments. First, in some middle-income countries, vital registration is incomplete. Wang and colleagues have used death distribution methods[43] to assess completeness in all countries with vital registration; we have used this information to correct upwards incomplete registration. Second, a key aspect of the GBD is to redistribute deaths that are assigned to immediate or intermediate causes of death rather than underlying causes of death (garbage codes).[74] In addition to garbage codes, because HIV was not included in the International Classification of Diseases (ICD)-9 until 1986 and not implemented in many countries until later, deaths were often assigned to other codes such as graft versus host disease or Kaposi's sarcoma. Third, in some places, because of stigma or misdiagnosis, HIV deaths can be assigned to other underlying causes of death such as tuberculosis, endocrine disorders, meningitis, or encephalitis. Birnbaum and colleagues developed a method to identify these misclassified deaths.[75] We applied this method to all

countries. Misclassification of HIV deaths in Birnbaum and colleagues' method is based on fulfilling three criteria. First, the temporal trend for a cause should coincide with the HIV epidemic. Second, the pattern of the relative death rate by age should shift towards the ages of 15–49 years during the epidemic years. Third, the temporal and age-pattern shifts cannot be explained by other known epidemiological trends. Applying these methods, we transfer deaths from 47 causes in 52 countries. Figure 3 shows the number of HIV deaths directly coded to HIV, the number of deaths re-assigned to HIV from garbage code redistribution, and the number of deaths from the application of the HIV misclassification method for Thailand and Russia. The height of the bar is the final number of deaths in each age group.

For countries with corrected vital registration, we imputed missing years of data to generate a complete time series for HIV from the estimated start year of the epidemic using spatial-temporal Gaussian Process Regression (ST-GPR).[43,76,77] ST-GPR using a linear mean function and a Matern covariance has been widely used for time-series estimation in global health descriptive studies such as for tobacco prevalence, obesity prevalence, or child and adult mortality. To generate an internally consistent set of incidence, prevalence, and death curves with uncertainty, we used the observed HIV death numbers to calibrate the modified Spectrum models. First, we started with a modified Spectrum model constructed based on the analysis of high-risk groups— where no high-risk group analysis was available we used a regional default model. Second, for each of the 1000 draws from this model, we modified the incidence at time t by the ratio of observed deaths to modified spectrum deaths at time t $+\lambda$, where λ is the lag between incidence and death. We drew from a distribution of lags of 10-15 years to generate 6000 different adjusted incidence curves. For incidence for the last λ time periods, we drew a random weight between 0 and 1 from a uniform distribution and used it to calculate a weighted average of adjusted incidence in year $t + \lambda$ and unadjusted incidence multiplied by the deaths ratio in year $t + \lambda$. Using these modified incidence curves, we reran the modified Spectrum generating 6000 possible epidemic curves. As a final step, for each of these 6000 we computed the mean squared error of predicted deaths compared with observed deaths. The 1000 curves with the lowest mean squared error were selected as the final set for analysis. Figure 4 shows the results of this process for Panama and the comparison with the UNAIDS high-risk group analysis; the corrected vital registration data suggest a much smaller epidemic with different timing.

Concentrated epidemics with high-risk group analysis and insufficient or no cause of death data—There were 17 countries with concentrated epidemics where we had no or insufficient vital registration or verbal autopsy data to inform our cause of death analysis. For these countries we ran modified Spectrum to output 1000 draws of incidence, randomly selecting 1000 time series of the death ratios generated in the process described above, and multiplying each draw of incidence by the selected set of ratios. We selected incidence adjustments only for countries with a cumulative number of HIV deaths greater than 5000 to avoid exaggerated stochastic variation in the ratios. We then derived estimates of mortality by running the adjusted incidence curves back through Spectrum. By using random draws across these countries, the average correction and uncertainty in this correction is propagated into the corrections for these countries with little or no data.

Concentrated epidemics with no high-risk group analysis and no cause of death data—For 13 countries (Andorra, United Arab Emirates, Iraq, Federated States of Micronesia, Libya, Marshall Islands, State of Palestine, Solomon Islands, Timor-Leste, Vanuatu, Samoa, Tonga, and North Korea) no analysis of high-risk groups has been undertaken and no cause of death data are available. For these countries, we picked regional or neighbouring countries to approximate the death rate. We used these approximate death rates to fit a Spectrum model. In all these cases, the number of estimated deaths was less than 250 per year.

Comparisons to prevalence survey data—As a form of empirical validation, we compared our final estimated prevalence with national population-based surveys collected through the Demographic and Health Surveys, AIDS Indicator Surveys, the 2005–2006 Indian National Family Health Survey, and the 2012 South African National HIV Prevalence, Incidence, and Behaviour Survey.[78-80] In total, we extracted data from 46 surveys in 35 countries between 2001 and 2012. These surveys had response rates for HIV testing ranging from 63% in male respondents in Malawi in 2004 to 98% for both sexes in Rwanda in 2011; median response rate was 85%. These comparisons are made for adults aged 15–49 by sex and 5-year age groups. We tested for significant differences in means ($p<0{\cdot}05$) for each estimate and compared the distribution of survey estimates to GBD and UNAIDS via ordinary least squares (OLS) regression with robust standard errors to account for heteroscedasticity.

HIV intervention counterfactual scenario—Spectrum uses as inputs the numbers reported by governments of individuals receiving PMTCT, co-trimoxazole, and child and adult ART. To help understand the role of interventions including ART, PMTCT, and co-trimoxazole prophylaxis, we have rerun the final 1000 modified Spectrum models for each country using a no intervention counterfactual scenario. We turn all HIV-related interventions to zero including ART, PMTCT, and co-trimoxazole prophylaxis for all years. We compared the number of deaths and person-years lived each year from the base case to this counterfactual to assess the changes due to intervention. Using published results on DAH for HIV, we computed the ratio of DAH to years of life saved.

Tuberculosis

For the GBD 2010, we estimated tuberculosis mortality and then estimated population incidence through mixed effects regression as a function of tuberculosis mortality, case-notifications, and an indicator variable for health system access used as a proxy for completeness of registration. For GBD 2013, we have shifted to using all available data for different outcomes and simultaneously estimating incidence, remission, excess mortality, prevalence, and cause-specific mortality using the GBD Bayesian meta-regression environment, DisMod-MR 2.0 (appendix p 11). There are four potential sources of information to estimate national levels and trends for tuberculosis in a country: annual case notifications, expert judgment on the case-detection rate, prevalence surveys, and cause of death data. Additionally, to facilitate convergence of the meta-regression, estimated excess mortality and remission rates have been used. The approach is predicated on the principle that incidence, prevalence, and mortality might be measured imperfectly and that a statistical

Europe PMC Funders Author Manuscripts

triangulation of all the sources for a country will provide a more robust assessment. Our meta-regression analysis was done for all forms of tuberculosis. As a final step we estimated incidence, prevalence, and death in individuals who are HIV-positive and those who are HIV-negative. We explain in more detail the preparation of each of these sources.

Adjusted case notifications and incidence—Case definitions for tuberculosis since 1995 have been standardised by WHO and widely applied. Countries have varied however in the completeness of reporting for younger age-groups and some countries have reported only pulmonary smear-positive cases for selected years. We use the age and sex-specific notifications in our analysis and impute the missing age-groups for three forms of tuberculosis notifications (pulmonary smear-positive, pulmonary smear-negative, or extra-pulmonary) in two steps. First, for each country-sex category with missing age-groups in some years, we imputed the missing values by regressing the log of the case notification rate on dummy variables for 5-year age-groups and random effects on year using all the data for a country over the interval 1990–2013.

Second, we estimated the relation between all forms of tuberculosis and smear-positive tuberculosis and the relation between all forms and bacteriologically positive tuberculosis. Using country-years with complete notifications (all three forms), we used a compositional analysis model to simultaneously estimate the fraction of cases due to all three forms as a function of dummy variables for 5-year age-groups and sex and the smear-positive tuberculosis rate. This regression was used to impute missing values for pulmonary smear-negative and extra-pulmonary cases. Because of substantial variation in the diagnostic rates for extrapulmonary tuberculosis and the potential for misclassification, we used the predicted values for extrapulmonary cases for all countries from the regression.

At the country-level several smear-unknown and relapsed cases are recorded that are not captured in the age-specific and sex-specific notifications. We used the relation between these forms and pulmonary smear-positive cases in the country-level data to inflate the adjusted age-specific and sex-specific pulmonary smear-positive notifications used in our analysis. Case notifications, however, do not capture all true incidence cases in the population. Case notifications can be incomplete because some cases are not diagnosed and some diagnosed cases are not reported to the national tuberculosis programme. Population-based incidence studies for tuberculosis based on active surveillance are rare and have not been done at the national level.[81] In the absence of direct measurement of true incidence, the case detection rate must be approximated. Since 2008, WHO has been consulting with national tuberculosis programme managers in 96 countries to collect expert opinion on the case detection rate including some notion of subjective uncertainty.[82] For the remaining countries, case-detection estimates are based on the judgment of WHO staff. We divided adjusted case notifications by the estimated case detection rate to generate the incidence inputs used for DisMod-MR 2.0. We expanded the subjective uncertainty intervals reported so they are at a minimum plus or minus 20% from the estimated values or for values less than 20% we assume the standard error is half the midpoint estimate.

Tuberculosis prevalence surveys—Prevalence surveys have been periodically undertaken in a few countries such as South Korea and China. WHO standardised the

protocol and the Global Fund to Fight AIDS, Tuberculosis and Malaria has helped fund 24 surveys in 21 countries between 2002 and 2013 with 12 additional surveys in eight new countries planned.[82-84] Because the prevalence rates for tuberculosis are often comparatively low (eg, 200 per 100 000 population), prevalence surveys need to be large to provide breakdowns by age and sex. On the basis of the literature and country reports, we have identified 27 national and 24 subnational prevalence surveys in 24 countries spanning the time period 1985–2013. We have included in our analysis, surveys reporting on pulmonary smear-positive tuberculosis and bacteriologically positive tuberculosis. We included in the Bayesian meta-regression study level dummy variables for the different measured outcomes with the reference category being bacteriologically positive. We allowed for non-sampling variance for sub-national surveys to be larger which effectively down-weighted their importance for the estimation in a given country. Because mortality and incidence data are for all forms of tuberculosis, we adjusted prevalence surveys to account for extra pulmonary cases using the same factors used in the adjustment of case notification data.

Mortality—We used 2731 country-years of nationally representative vital registration data and 166 site-years of verbal autopsy data to estimate tuberculosis mortality. Vital registration data were adjusted for garbage coding following GBD algorithms[74,85] and misclassified HIV deaths described above. We modelled deaths by age and sex for tuberculosis using the Cause of Death Ensemble modelling (CODEm) approach.[76] CODEm has been extensively used in global cause of death analyses.[74] Using CODEm, we tested a wide range of potential models and used out-of-sample predictive validity to select the best individual models and the best ensemble of these models. The appendix (pp 152–186) shows details on the application of CODEm to tuberculosis. We ran CODEm separately for male and female individuals. The final ensemble models selected had a root-mean squared error of the log of the age-specific death rate of 0·29 in-sample and 0·63 out-of-sample for males and 0·70 in-sample and 1·05 out-of-sample for females. In the out-of-sample predictive validity testing (cross-validation), the coverage of the 95% data prediction uncertainty interval was 93% and 91% for males and females, respectively.

CODEm results are largely informed by ICD-coded data, which by definition exclude tuberculosis mortality in HIV-positive individuals. The overall Bayesian model, however, is for all forms of tuberculosis in HIV-negative and HIV-positive people because prevalence data rarely distinguish HIV status. We estimated the fraction of HIV deaths due to tuberculosis-HIV and added these to tuberculosis mortality in HIV-negative people. The model for the fraction of tuberculosis-HIV mortality was based on 1022 country-years of data when cause of death data for tuberculosis-HIV and HIV overall were available. We estimated the relation between the logit-transformed fraction of HIV deaths due to tuberculosis-HIV and the log-transformed tuberculosis death rate, a dummy variable for sex, year, and country random effects. We used this regression to predict the fraction of HIV deaths due to tuberculosis-HIV in all countries.

Remission and excess mortality estimates—To help inform the model, we generated a Bayesian prior for remission by examining the ratio of incidence to prevalence in the

country-years where prevalence surveys have been undertaken. We used a simple regression with random effects to generate priors for countries with surveys and those without. Cause-specific mortality estimates inform estimates of prevalence through excess mortality in DisMod-MR 2.0. To provide the model with the range of age-specific and sex-specific excess mortality hazards associated with tuberculosis we analysed historical data where we had both tuberculosis mortality data and incidence data that were believed to be nearly complete. For this analysis, we used the WHO case notifications from 1980 onwards with the supplement of age-sex-specific case notifications back to the 1950s for Australia, Canada, the UK, USA, Japan, and Germany. Case notification data were combined with tuberculosis deaths recorded in the vital registration systems to generate 743 country-year observations from 70 countries that could be used to inform our analysis. We estimated the relation between incidence and mortality for each sex, by regressing the logit-transformed ratio of incidence to mortality against age, lag-distributed income per head, and country random effects. The addition of HIV prevalence off-ART to the regression gave inconsistent coefficients between females and males and was not included in the final model. We estimated the relation between incidence and prevalence as a function of lag-distributed income per head with country random effects. We transformed predicted death to incidence ratios and incidence to prevalence ratios into estimates of excess mortality and remission using the mathematical relations between them (appendix p 7).

DisMod-MR 2.0—For each country we included in the DisMod-MR 2.0 estimation the adjusted case notifications, prevalence survey data if available, estimated excess mortality hazard by age and sex, estimated remission, and the tuberculosis-HIV adjusted cause-specific mortality estimates from our CODEm model. DisMod-MR 2.0 provides internally consistent estimates for 1990, 1995, 2000, 2005, 2010, and 2013 for 188 countries of incidence, remission, excess mortality, prevalence, and cause-specific mortality using all forms of data or priors in the estimation. Figure 5 shows the internally consistent fit for Kenya in 2013. For intervening years, we interpolated rates.

Estimating tuberculosis incidence, prevalence, and death in individuals who are HIV-positive—We used tuberculosis all-forms estimates from DisMod-MR 2.0 to estimate incidence and prevalence in HIV-positive people using a relative risk approach. We reviewed the literature using the search terms "incidence", "risk ratio", "HIV", "tuberculosis", and "antiretroviral therapy" and used meta-regression to estimate a relative risk of tuberculosis incidence in HIV-positive individuals in the absence of ARTs based on seven studies[86-92] of 8·7 (95% CI 5·9–11·7). Findings from previous studies show that the relative risk of tuberculosis incidence is a function of CD4 count and ART; to parse out the increasing risk ratios of tuberculosis by decreasing CD4 count and the decreasing risk ratio on ART we used data from the Badri and colleagues' study.[93] The relative risks we calculated from this analysis were 15·7 (10·6–21·1) for a CD4 cell count less than 200, 10·8 (7·3–14·5) for a cell count of 200–350, 3·2 (2·2–4·3) for a count greater than 350, and 1·7 (1·2–2·3) for the on-ART category. We computed population-attributable fractions for each category using the outputs of Spectrum above. For prevalence, we assumed that each category of incident tuberculosis cases in HIV-positive individuals has the same duration. Tuberculosis-HIV mortality was estimated as described above.

Europe PMC Funders Author Manuscripts

Malaria

Murray and colleagues developed estimates of mortality and incidence for malaria for the GBD 2010.[74,94] They estimated malaria mortality using vital registration and verbal autopsy data analysed using CODEm. Published community incidence studies were meta-analysed to generate a model of incidence as a function of mortality, age, sex and region. We have modified this method for the GBD 2013 update. Much debate emerged since the publication of that analysis on the validity of verbal autopsy for adult malaria deaths.[95-100] For the GBD 2013, we undertook a systematic review of the literature on the validity of verbal autopsy for malaria. Our inclusion criteria were validation studies that used physician-certified verbal autopsies, reported both sensitivity and specificity for malaria, and had hospital diagnosis as the gold standard. However, the quality of the gold standards used in these studies was variable, and in some of them malaria cases were not confirmed with a blood smear or did not use a case definition with a threshold of parasitaemia. We identified seven studies.[101-107] We first tested in a meta-regression if there was any statistically significant difference between studies with and without parasitaemia confirmation and identified none. We meta-analysed these studies to estimate sensitivity and specificity, separately for children and adults. Forest plots for adult and children are shown in figure 6. As a sensitivity analysis, we used this correction but it leads to substantially larger numbers of estimated deaths in adults from malaria (appendix p 21). We have chosen not to correct the data for the main results of this paper because it would adjust deaths in adults upwards which is contrary to expert opinion in the literature.

In view of the fact that we have not applied the sensitivity and specificity corrections, we have instead modified the redistribution of garbage codes such as fever of unknown origin or ill-defined deaths, so that we do not redistribute garbage codes to malaria in adults. We have also updated all the times-series covariates tested in the models: rainfall, health-system access, antimalarial drug resistance weighted by drug use, ITN coverage, indoor residual spraying coverage, income per head, and educational attainment. We have also included in the model the 2010 *P falciparum* parasite rate (PfPR) map from the Malaria Atlas Project.[108] A coherent analysis of PfPR overtime is underway but was not available for this analysis (see the appendix pp 223–40 for details on the CODEm model analysis). As in the Murray and colleagues study, we developed separate models for sub-Saharan Africa and outside of Africa (with the exception of South Africa, which was modelled with countries outside of Africa, given the low malaria endemicity), age under 5 years and 5 years or higher, and males and females.

For countries that have only or mainly *Plasmodium vivax* transmission we used the number of deaths by year and age from vital registration data as a simple predictor of malaria mortality using a negative binomial regression model.

We estimated malaria cases separately for three sets of countries, which were divided on the basis of the availability and quality of malaria incidence data (see appendix pp 241–42 for the list of countries). The first group contained countries with unavailable or unreliable malaria case reporting systems. We estimated malaria incidence in these countries using a mortality-incidence model, in which we predicted malaria incidence by regressing the log-