The American Journal of Human Genetics, Volume *90*

Supplemental Data

A Subset-Based Approach Improves Power and Interpretation for the Combined-Analysis of Genetic Association Studies of Heterogeneous Traits

Samsiddhi Bhattacharjee, Preetha Rajaraman, Kevin B. Jacobs, William A. Wheeler, Beatrice S. Melin, Patricia Hartge, GliomaScan Consortium, Meredith Yeager, Charles C. Chung, Stephen J. Chanock, and Nilanjan Chatterjee

Figure 1

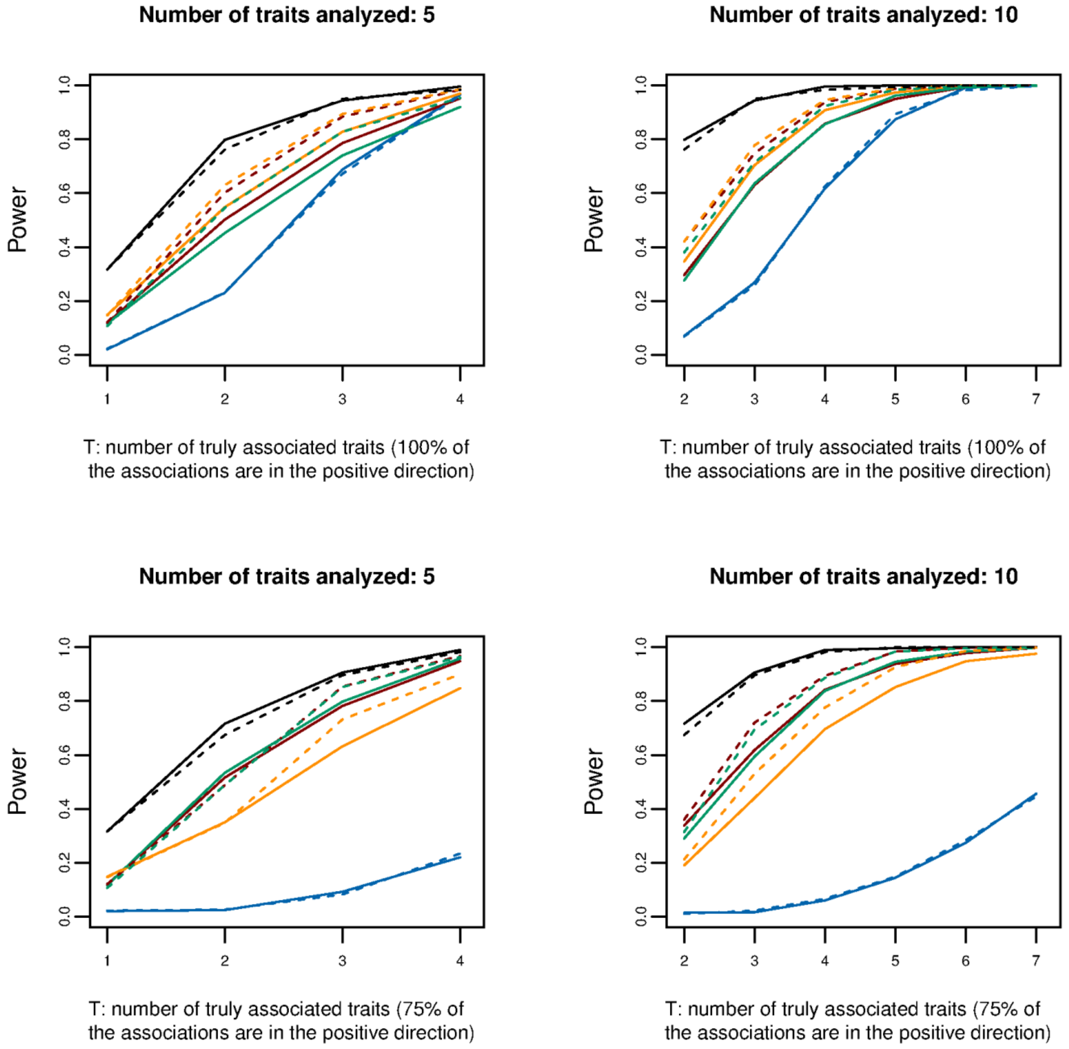

Figure S1. Simulation-Based Power Comparison of Alternative Methods for Detecting an Overall Association.

In each simulation, a total of K=5 or K=10 distinct traits are analyzed, each with 2000 cases and 2000 controls. A variant with MAF=0.3 is assumed to be associated with a subset of the traits (the number of such traits is shown on the x-axis). The solid and dashed lines represent, respectively, power when all the OR-s for associated traits are fixed at a single value (1.15) (as in Figure 1) and power when the OR-s are allowed to vary around a fixed mean (1.15) witihin a given range (1.05-1.25; see Supplemental Methods, section 6 for details). The upper panel assumes that all of the associations are in the same direction and the lower panel assumes that 75% of the associations are positive and 25% are negative. In addition to two-sided (green lines) and one-sided (orange lines) subset-based tests, power curves are shown for standard meta-analysis (blue lines), Fisher's combined p-value method – a multi degree-of-freedom chi-square test (maroon line) and a "gold-standard" test (black lines) that assumes that the subset of the traits that are truly associated is known a priori. All powers are shown at a level of 0.001.

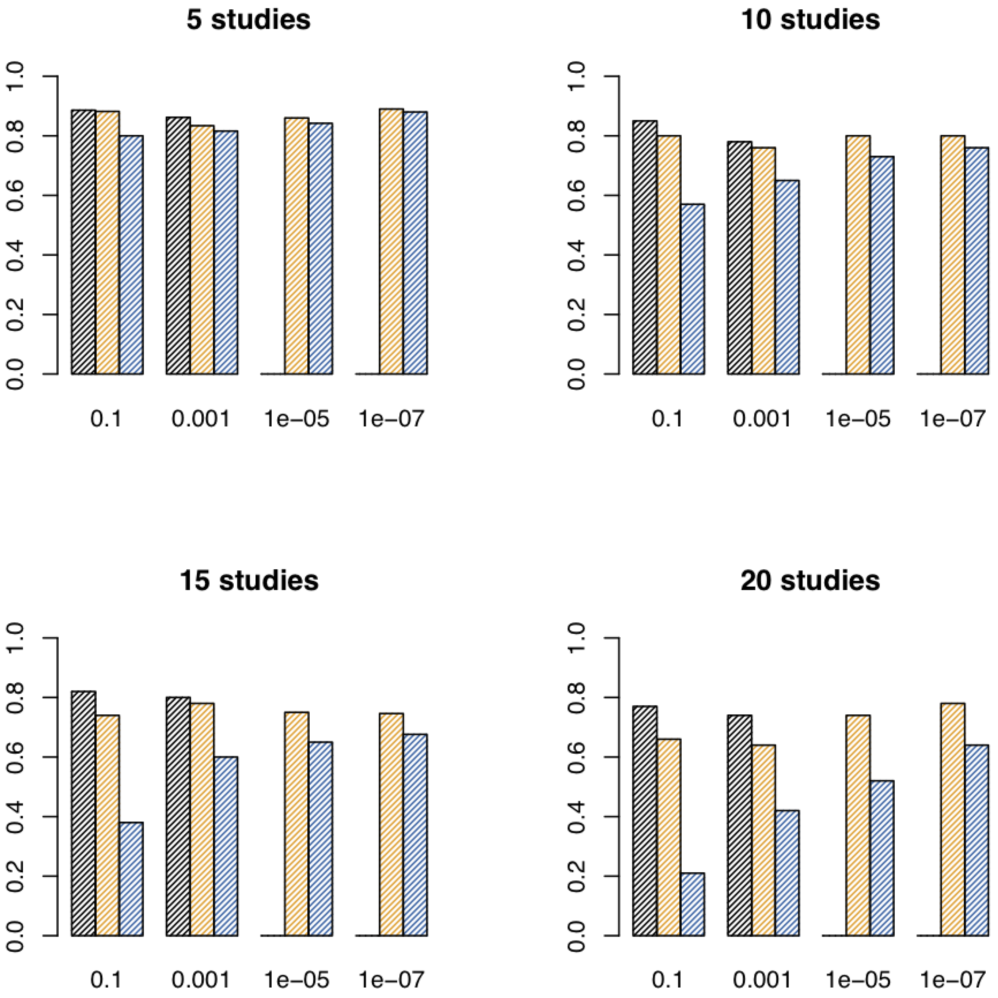

Figure S2. Simulation-Based Power Comparison of One-Sided Subset Search with Three Different p Value Estimation Procedures: Bonferroni (blue bars), DLM (orange bars) and Parametric Bootstrap (black bars)

The bootstrap method, which resamples effect size estimates for each study from a normal distribution with mean zero (assuming the null hypothesis of no association) and standard deviation equal to original standard error estimates, gives an “exact” method for computing p-values. In each simulation a total of K (5, 10,15 or 20) distinct traits are analyzed, each with equal number of cases and controls. A variant with MAF=0.3 is assumed to be associated with a subset of size T (3, 6, or 10 or 12) traits with a fixed odds ratio of 1.15. Sample sizes were chosen such that the theoretical power of the gold-standard test (not shown) is close to 95%. The total sample sizes for all traits N and that for the subset of associated traits N_T were fixed as follows: 1) $(N, N_T) =$ (5000, 3000), 2) $(N, N_T) =$ (10000, 6000), 3) $(N, N_T) =$ (15000, 9000) and 4) $(N, N_T) =$ (20000, 12000) for power-comparison at levels $\alpha =$ 0.1, 0.001, 10^{-5} and 10^{-7} respectively. In each case, the DLM method performed superior to Bonferroni. The power of the Bootstrap method was computed only at the levels $\alpha =$ 0.1 and 0.001 (due to compuational limitations), where it performed slightly better than DLM but considerably better than Bonferroni.

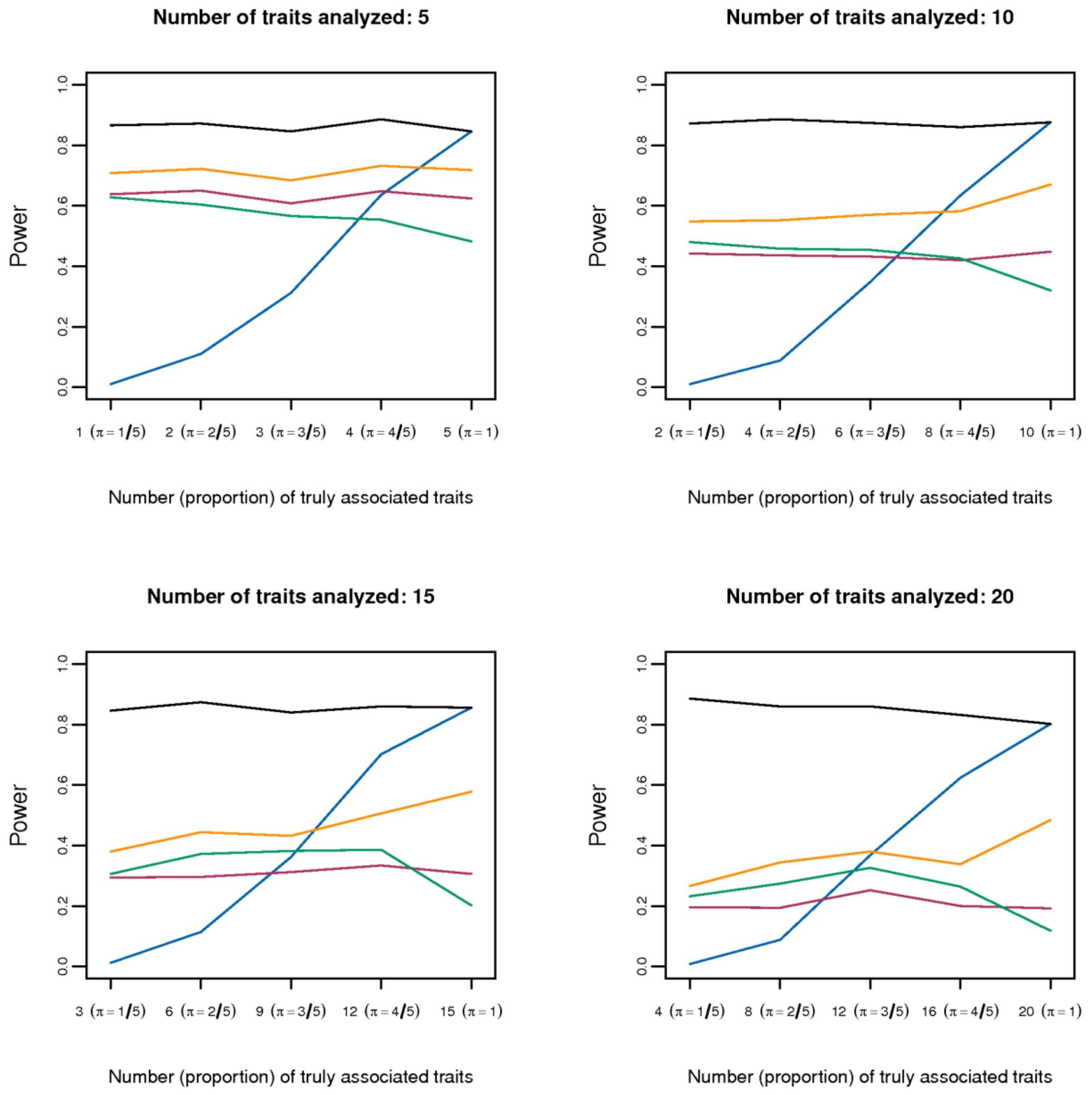

Figure S3. Simulation-Based Power Comparison for Combined Analysis of Heterogeneous Traits

In each simulation, a total of K (K=5, 10, 15 and 20) independent case-control studies are analyzed each with equal number of cases and controls. A variant with MAF=0.3 is assumed to be associated with a subset of the traits (non-null traits) with a fixed odds-ratio of 1.1. The number of non-null traits (T) and the sample sizes for individual studies are varied such that the total number of cases/controls for associated traits is held fixed at M=10000, but the ratio π =M/N, the fraction of total sample size that contains a true association signal varies (π = 1/5, 2/5, 3/5, 4/5, or 1). In addition to the two-sided (green line) and one-sided (orange line) subset-based tests, power curves are also shown for overall meta-analysis (blue line), Fisher's combined p-value method – a multi degree-of-freedom chi-square test (maroon line), and that for a "gold-standard" meta-analysis (black line) that assumes the subset of traits that are truly associated with the given SNP is known a priori. All powers are shown at a level of 10^{-7}.

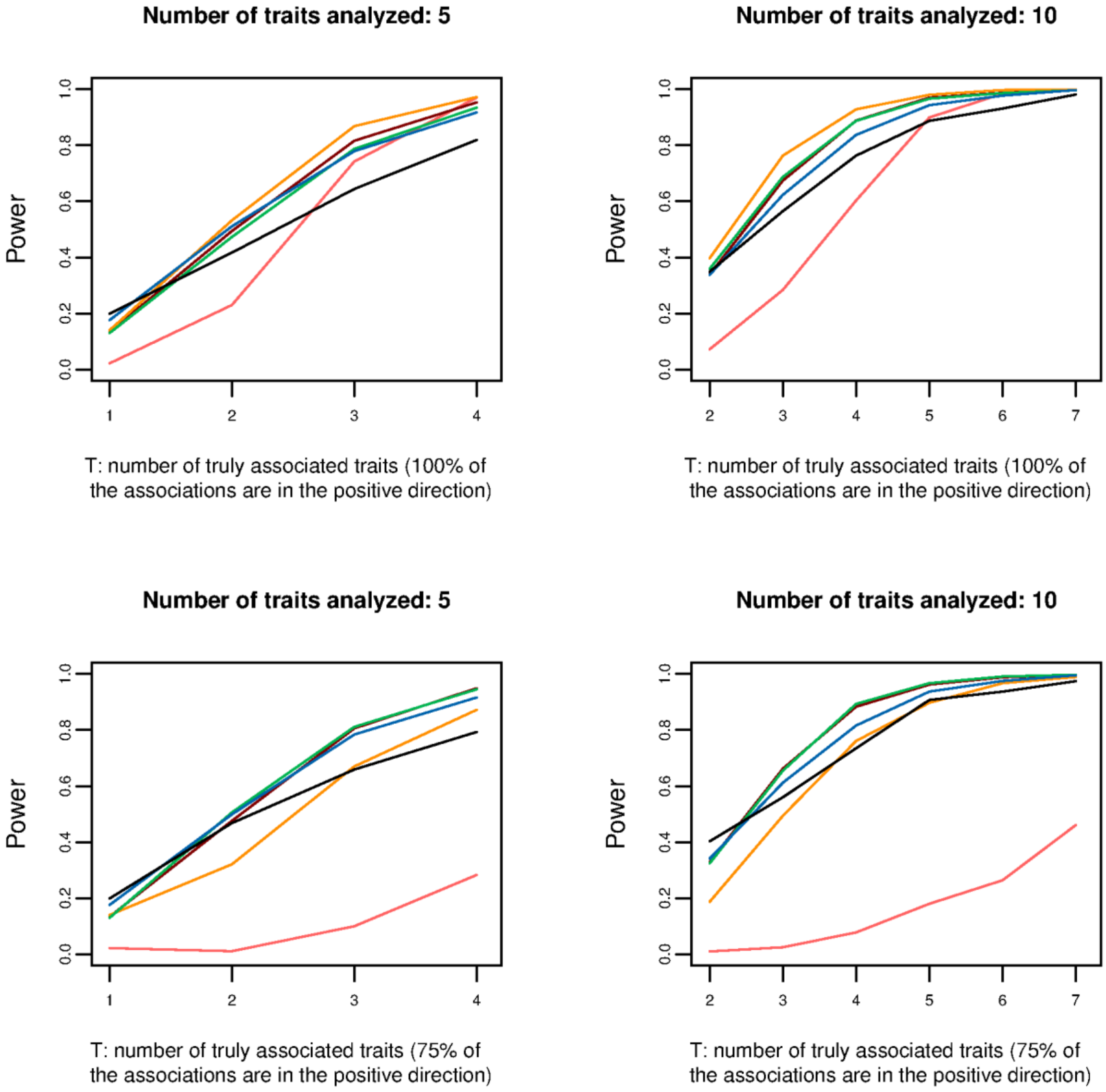

Figure S4. Simulation-Based Power Comparison of Alternative Methods for Detecting an Overall Association with Homogeneous Effect Sizes

In each simulation, a total of K=5 or K=10 distinct traits are analyzed, each with 2000 cases and 2000 controls. A variant with MAF=0.3 is assumed to be associated with a subset of the traits (the number of such traits is shown on the x-axis). The OR-s for associated traits are fixed at a single value (1.15) (as in Figure 1). The upper panel assumes that all of the associations are in the same direction and the lower panel assumes that 75% of the associations are positive and 25% are negative. In addition to two-sided (green lines) and one-sided (orange lines) subset-based tests, power curves are shown for standard meta-analysis (red lines), Fisher’s combined p-value method – a multi degree-of-freedom chi-square test (maroon line), the Adaptive Rank Trunkated Product (ARTP) test with K truncation points (black line) and the Adaptively Weighted (AW) statistic (blue line). All powers were calculated using empirical (simulation-based) cutoff at level 0.001.

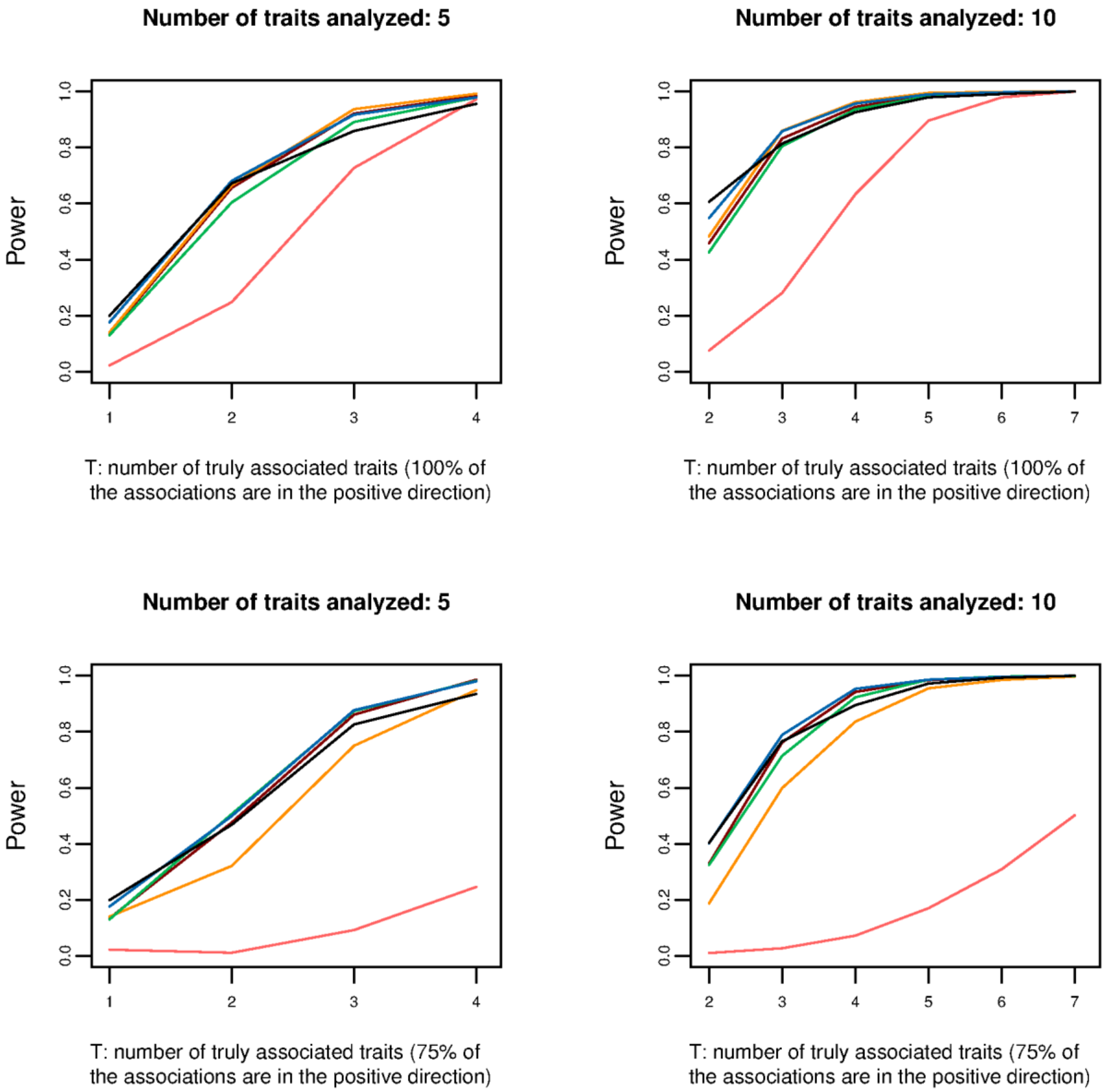

Figure S5. Simulation-Based Power Comparison of Alternative Methods for Detecting an Overall Association with Heterogeneous Effect Sizes

In each simulation, a total of K=5 or K=10 distinct traits are analyzed, each with 2000 cases and 2000 controls. A variant with MAF=0.3 is assumed to be associated with a subset of the traits (the number of such traits is shown on the x-axis). The OR-s for associated traits are allowed to vary around a fixed mean (1.15) witihin a given range (1.05-1.25; see Supplemental Methods, section 6 for details). The upper panel assumes that all of the associations are in the same direction and the lower panel assumes that 75% of the associations are positive and 25% are negative. In addition to two-sided (green lines) and one-sided (orange lines) subset-based tests, power curves are shown for standard meta-analysis (red lines), Fisher’s combined p-value method – a multi degree-of-freedom chi-square test (maroon line), the Adaptive Rank Trunkated Product (ARTP) test with K truncation points (black line) and the Adaptively Weighted (AW) statistic (blue line). All powers were calculated using empirical (simulation-based) cutoff at level 0.001.

Table 1. Performance of the Subset-Based Test for Detection of the Truly Associated Subset of Traits in Presence of Heterogeneity of Odds Ratios

(K, T1, T2)	**Sensitivity (True Positive Probability)**		**Specificity (True Negative Probability)**	
A. 5 studies	One-sided	Two-sided	One-sided	Two-sided
100% positive				
(5, 1, 0)	0.920	0.986	0.835	0.500
(5, 2, 0)	0.754	0.773	0.919	0.541
(5, 3, 0)	0.745	0.763	0.939	0.498
(5, 4, 0)	0.748	0.764	0.948	0.440
75% positive				
(5, 1, 1)	0.502	0.982	0.883	0.721
(5, 2, 1)	0.495	0.829	0.931	0.775
(5, 3, 1)	0.547	0.807	0.946	0.800
B. 10 studies				
100% positive				
(10, 2, 0)	0.765	0.785	0.909	0.590
(10, 3, 0)	0.762	0.773	0.929	0.614
(10, 4, 0)	0.762	0.772	0.942	0.616
(10, 5, 0)	0.764	0.780	0.945	0.595
(10, 6, 0)	0.752	0.766	0.950	0.585
(10, 7, 0)	0.753	0.770	0.955	0.549
75% positive				
(10, 1, 1)	0.500	0.978	0.855	0.675
(10, 2, 1)	0.502	0.838	0.907	0.739
(10, 3, 1)	0.558	0.818	0.930	0.768
(10, 3, 2)	0.454	0.772	0.944	0.839
(10, 4, 2)	0.501	0.760	0.947	0.854
(10, 5, 2)	0.539	0.765	0.951	0.868

In each simulation, a total of K=5 or K=10 distinct traits are analyzed, each with 2000 cases and 2000 controls. A variant with MAF=0.3 is assumed to be associated with a subset of size T (< K) traits with odds ratios varying within the range 1.05-1.25 with a mean value of 1.15 (see Methods for Details). For each K (i.e., 5 or 10 traits), Panel A assumes that all of the associations are in the same direction and Panel B assumes that 75% of the associations are positive and 25% are negative. Two measures of performance are shown; (1) sensitivity: the average proportion of associated traits detected and (2) specifity: the average proportion of null traits discarded.

§ K = Total number of traits analyzed (5 or 10).
T_1 = Number of traits that are truly associated in the positive direction.
T_2 = Number of traits that are truly associated in the negative direction.

Table S2. The Number of Cases and Controls and Their Overlaps[†] in the NCI GWAS Studies of Six Cancer Sites

Site	BLADDER	BREAST	KIDNEY	LUNG	PANCREAS	PROSTATE
Cases Shared:						
BLADDER	3574	0	3	47	4	37
BREAST	0	1726	0	0	0	0
KIDNEY	3	0	2877	9	3	6
LUNG	47	0	9	5805	2	99
PANCREAS	4	0	3	2	3954	14
PROSTATE	37	0	6	99	14	3537
Controls Shared:						
BLADDER	6200	63	8	2960	227	3609
BREAST	63	1472	0	189	0	0
KIDNEY	8	0	3754	5	0	4
LUNG	2960	189	5	5315	4	2378
PANCREAS	227	0	0	4	4097	46
PROSTATE	3609	0	4	2378	46	5053

There were a total of 21,473 cases and 25,891 controls.

[†] Numbers along the diagonals equal total sample size for that study. Numbers outside the diagonal represent shared cases (upper panel) and shared controls (lower panel).

Table S3. Details of the 18 GliomaScan Studies Included in the Glioma Subtype Analysis Example

Study	**Study Acronym**	**Study Type**	**Control Selection**	**Study Period (recruitment)**	**Location**	**Cases**	**Controls**	**Mean age at diagnosis, cases (yrs)**
Agricultural Health Study	AHS	Cohort	Frequency-matched 2:1 by year of birth, sex, race	1993-1997	USA (IA, NC)	18	35	57.2
Alpha-Tocopherol Beta-Carotene Cancer Prevention Study	ATBC	Cohort	Glioma-free controls with previous GWAS data	1985-1993	Finland	37	1,270	69.3
Campaign Against Cancer and Stroke; Cancer and Heart Disease	§CLUE -I & II	Cohort	Glioma-free controls with previous GWAS data	1974-1989	USA (MD)	36	71	65.0
Cancer Prevention Study II Nutrition Cohort	CPS-II	Cohort	Glioma-free controls with previous GWAS data	1992	USA (21 states)	54	98	73.0
Gliogene (Sweden)	GLIOGENE	Family	Glioma-free controls from Sweden	2004 - present	Sweden	401	712	54.2
Health Professionals Follow-up Study	HPFS	Cohort	Glioma-free controls with previous GWAS data	1986 - ongoing	USA	26	52	68.1
Interphone Study (Sweden, Denmark)	INT-SD	Case-control	Matched on year of birth, sex, study region	2000-2004	Sweden, Denmark	277	381	49.5
Melbourne Collaborative Cohort Study	MCCS	Cohort	Glioma-free controls with previous GWAS data	1990-1994	Australia	40	75	68.0
Multiethnic Cohort	MEC	Cohort	Matched on age, sex, race	Age 45-75 in 1993	US	2	6	73.5
National Cancer Institute Adult Brain Study	NCI-BTS	Case-control	Frequency-matched 2:1 by hospital, age, sex, race, residential distance from hospital	1994-1998	USA (AZ, MA, PA)	322	385	51.5

Upper Midwest Health Study - National Institute for Occupational Safety and Health	[†]NIOSH-UMHS	Case-control	Population-based controls frequency matched 1:5:1 by age, sex, state	1995-1997	USA (IA, MI, MN, WI)	300	542	48.7
New York University - Women's Health Study	NYUWHS	Cohort	Glioma-free controls with previous GWAS data	1985 - 1991	USA (NY)	5	12	65.4
Northern Sweden Health and Disease Study	NSHDS	Cohort	Matched on age, sex, race	1985 - ongoing	Northern Sweden	111	222	52.1
Nurses' Health Study I and II	NHS	Cohort	Glioma-free controls with previous GWAS data	1976, 1989	USA (several states)	38	83	66.9
Physician's Health Study I and II	PHS	Cohort	Glioma-free controls with previous GWAS data	1982 – 1984, 1997 – 2001	USA (several states)	16	54	71.1
Prostate, Lung, Colorectal and Ovarian Screening Trial	PLCO	Cohort	Glioma-free controls with previous GWAS data	1992-2001	USA (several states)	132	855	71.2
Vitamins and Lifestyle	VITAL	Cohort	Matched on age, sex, race, time to diagnosis	2000-2002	USA (WA)	33	71	68.5
Women’s Health Study	WHS	Cohort	Glioma-free controls with previous GWAS data	1991-2009	USA (several states)	8	31	58.3

There were a total of 1,856 cases and 4,955 controls.

[§] Cancer incidence data for the CLUE study have been provided by the Maryland Cancer Registry. We acknowledge the State of Maryland, the Maryland Cigarette Restitution Fund, and the National Program of Cancer Registries of the Centers for Disease Control and Prevention for the funds that helped support the availability of the cancer registry data.

[†] The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

Table 4. Simulation-Based Estimates of Type I Error for Combined Analysis of Heterogenous Traits

(N, K)[§]	Meta Analysis	Subset-based	
		One-sided	Two-sided
Level 0.05			
(2000, 5)	0.0480	0.0570	0.0430
(2000, 10)	0.0420	0.0310	0.0300
Level 0.01			
(2000, 5)	0.0106	0.0096	0.0078
(2000, 10)	0.0114	0.0088	0.0064
Level 0.001			
(2000, 5)	0.0010	0.0010	0.0010
(2000, 10)	0.0011	0.0009	0.0009

In each simulation, a total of K=5 or K=10 distinct traits are analyzed, each with N=2000 cases and N=2000 controls. A variant with MAF=0.3 unrelated to all the traits is tested for association in each case. Type I errors are shown for the standard meta-analysis and the proposed one-sided and two-sided methods based on subset search, at three different nominal significance levels.

[§] N = Number of cases from each study corresponding to a single trait (fixed at 2000).
K = Total number of traits analyzed (either 5 or 10).

Table S5. Simulation-Based Estimates of Type I Error for the Analysis of a Case-Control Study with Geterogeneous Disease Subtypes

(N, K, N_0)	**Overall Case-control**	**Subset-based Test**	
Level 0.05		Case-control	Case-complement
A. (2000, 7, 14000)	0.046	0.037	0.047
B. (2000, 7, 3000)	0.046	0.036	0.040
Level 0.01			
A. (2000, 7, 14000)	0.0114	0.0084	0.0102
B. (2000, 7, 3000)	0.0112	0.0094	0.0086
Level 0.001			
A. (2000, 7, 14000)	0.00108	0.00088	0.00082
B. (2000, 7, 3000)	0.00086	0.00076	0.00092

Each simulation includes N=2000 cases for each of K=7 subtypes. Rows labeled A and B correspond to designs with N_0=14000 and N_0=3000 shared controls respectively. A variant with MAF=0.3 unrelated to all disease subtypes is tested for association in each case. Type I errors are shown for an overall case-control analysis and the two alternative subset-based tests, namely “Case-control” and “Case-complement,” at three different nominal significance levels.

§ N = Number of cases from corresponding to each disease subtype (fixed at 2000).
K = Total number of traits analyzed (fixed at 7).
N_0 = Number of controls available (either 14000 or 3000).

The GliomaScan Consortium Investigators†

Demetrius Albanes[1], Ulrika Andersson[2], Laura Beane-Freeman[3], Christine D. Berg[4], Julie E. Buring[5], Mary Ann Butler[6], Tania Carreon[6], Helle Collatz Christensen[7], Maria Feychting[8], Susan M. Gapstur[9], J. Michael Gaziano[10,11], Graham G. Giles[12, 23], Goran Hallmans[13], Susan E. Hankinson[14], Roger Henriksson[15,27], Jane Hoppin[16], Ann W. Hsing[17], Peter D. Inskip[18], Christoffer Johansen[7], Laurence N. Kolonel[19], Roberta McKean-Cowdin[20], Dominique Michaud[21], Ulrike Peters[22,26], Mark P. Purdue[3], Avima M. Ruder[6], Howard D. Sesso[10], Gianluca Severi[12,23], Victoria L. Stevens[9], Kala Visvanathan[24,25], Zhaoming Wang[27], Emily White[22,26], Walter C. Willett[28], Anne Zeleniuch-Jacquotte[29]

[1] Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA
[2] Department of Radiation Sciences, Oncology, Umeå University, 90187 Umeå, Sweden
[3] Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA
[4] Early Detection Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, MD 20852, USA
[5] Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215, USA
[6] National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention, Cincinnati, Ohio, USA
[7] Institute of Cancer Epidemiology, Danish Cancer Society, DK-2100 Copenhagen, Denmark
[8] Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
[9] Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA
[10] Divisions of Preventive Medicine and Aging, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215, USA
[11] Massachusetts Veteran's Epidemiology, Research and Information Center, Geriatric Research Education and Clinical Center, VA Boston Healthcare System, Boston, Massachusetts, United States of America
[12] Cancer Epidemiology Centre, The Cancer Council of Victoria, Melbourne, Australia
[13] Department of Public Health and Clinical Medicine/Nutritional Research, Umea University
[14] Channing Laboratory, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215, USA
[15] Dept Oncology, Karolinska University hospital, Stockholm
[16] National Institute of Environmental Health Sciences, NIH, Research Triangle Park, NC, USA
[17] Infections and Immunoepidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA
[18] Radiation Epidemiology Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA
[19] Epidemiology Program, Cancer Research Center, University of Hawaii, Honolulu, Hawaii, USA
[20] Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA
[21] Department of Community Health, Brown University, Providence, RI, USA
[22] Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

[23] Centre for Molecular, Environmental, Genetic, and Analytic Epidemiology, The University of Melbourne, Melbourne, Australia
[24] Johns Hopkins Sidney Kimmel Comprehensive Cancer Center, Baltimore, MD, USA
[25] Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
[26] University of Washington, Department of Epidemiology; Seattle, WA, USA
[27] Core Genotyping Facility, National Cancer Institute, SAIC-Frederick, Inc., Gaithersburg, Maryland 20877, USA
[28] Departments of Nutrition and Epidemiology, Harvard School of Public Health, Boston, MA 02115, USA
[29] Division of Epidemiology, Department of Environmental Medicine, NYU School of Medicine, New York, NY 10016, USA

[†] Investigator names appear in alphabetical order.