An Early-life Conditions Scale, and Applications to Suicide and All-cause Mortality

Michael S. Hollingshaus, MS[1]

Ken R. Smith, PhD

Heidi A. Hanson, PhD

University of Utah

Acknowledgements:

- Funded by NIH R01 AG022095 Early Life Conditions, Survival, and Health: A Pedigree-based Population Study
- Huntsman Cancer Foundation

Draft Paper for the Annual Meeting of the *Population Association of America* at

San Diego CA, 2015

[1]PhD Candidate, Department of Sociology, University of Utah.
Graduate Research Assistant, Pedigree and Population Resource (Utah Population Database), Huntsman Cancer Institute.

380 S 1530 E Rm 301
Salt Lake City, UT 84112
email michael.hollingshaus@soc.utah.edu

ABSTRACT

The goal of this paper is to construct an early-life conditions scale that can be utilized to study later-life mortality and morbidity in historical demographic databases. We identified a sample of 536,850 individuals born between 1849 and 1972 in the Utah Population Database. Variables include advanced maternal and paternal age, large sibship size, early maternal and paternal death, high infant mortality rate in county of birth, late birth order, sibling died as an infant, sibling died as a child, low initial SES, being a twin or other multiple-birth sibling, out-of-wedlock birth, and short birth interval. Using a quarter sample of 134,213, we performed exploratory factor analysis. We present results from parallel analysis that suggest the variables consist of four factors. Factor patterns are also presented. After confirmatory factor analysis on a separate subsample, the final scale will be used to predict all-cause mortality and suicide hazards with Cox regression models.

INTRODUCTION

Several early life conditions (ELCS) have been shown to increase mortality risk in later life (*1*). Historical demographic databases are one source of data that have been utilized to examine the risks of ELCS (*2*). The purpose of this paper is to create a demographic ELCS scale that might be utilized by such databases, something we have not yet seen in the literature.

Benefits of this continuous scale would include a more finely-grained measure of ELCS for studying morbidity or mortality, potential standardization across studies, and the identification of possible commonalities between what may appear to be distinct stressors. This scale is also anticipated to serve as a prototype for a robust standardized demographic ELCS scale for use with the Utah Population Database (UPDB). Finally, it is intended to be utilized for a specific project examining ELCS correlated with later-life suicide risk.

For purposes of this version of the scale, we defined early-life as occurring before age 18, and later-life as afterwards. We identified thirteen putative early-life stressors that might reasonably be measured for a large group of cohorts in UPDB. For conceptual simplicity, each was operationalized as a dichotomous variable, measuring whether the stressor did or did not occur. These thirteen variables, hypothesized mechanisms, and example empirical literature for both all-cause mortality and suicide are displayed in Table 1.

Table 1. Thirteen early-life stressors, plausible mechanisms linking to later-life mortality, and example literature

Stressor	Plausible Mechanisms	Example Mortality Literature	
		All-Cause	**Suicide**
Advanced maternal age	Genetic mutation load	(*3-5*)	(*6, 7*)[a]
Advanced paternal age	Same as above	Same	Same
Large sibship size	Maternal depletion, low SES, early "scarring"	(*2, 8*)	(*7, 9*)
Early maternal death	Grief and allostatic load, loss of support, cumulative disadvantage	(*2, 10, 11*)	(*12, 13*)
Early paternal death	Same as above	Same	Same
High infant mortality rate for county of birth (exposure to infectious disease or cumulative disadvantage)	Early "scarring", inflammation, infections	(*14, 15*)	(*16-20*)
Late birth order	Maternal depletion, loss of bequests	(*21, 22*)	(*7, 23-25*)
Sibling died as infant (exposure to infectious disease or cumulative disadvantage)	Early "scarring", inflammation, infections	(*14, 15*)	(*16-18, 20*)
Sibling died as child	Grief and allostatic load, loss of support, cumulative disadvantage	(*26*)	(*27*)
Low initial SES	Cumulative disadvantage	(*28*)	(*7, 13, 25*)
Twin/multiple birth (possible indicator of low birth weight)	Biological programming	(*29, 30*)	(*16, 25, 31, 32*)[b]
Out-of-wedlock birth	Cumulative disadvantage	(*33*)	(*25*)
Short birth interval	Maternal depletion / biological programming	(*34*)	(*24, 35*)

Notes:

a. Research suggests advanced paternal age may increase risk for suicide, while increased maternal age may decrease risk.
b. Twins may have lower suicide risk than singletons.

Table adapted from a similar table in the application for NIH grant R01 AG022095 (PI Ken R. Smith).

METHOD

Data came from the Utah Population Database, a database for over seven million unique individuals from the genealogies of the founders of Utah as and their descendants. This dynamic database is updated annually using Utah birth, death, driver license, and health records. This includes over two million Utah birth certificates and around 800,000 death certificates. Access to the data is administered through the Utah Resource for Genetic and Epidemiologic Research (RGE). All research requires prior IRB and RGE approval (*36*).The confidentiality of individuals represented in these records is maintained based on agreements between RGE and the data contributors.

We used birth cohorts ranging from 1849 through 1972 (encompassing about 95% of the key demographic and pedigree cohorts in UPDB with relevant data). Our specific operationalizations for the thirteen variables are shown in Table 2. Cutoffs for dichotomous variables were empirically derived using the top or bottom 10^{th} percentile, as appropriate. Each individual had a known father and mother, at least one sibling, and survived to at least age 18. After imposing other data requirements necessary for variable construction, we obtained a sample of 536,850 individuals. Basic descriptive statistics for each stressor are also shown in Table 2.

Table 2. Thirteen early-life variables, our operationalizations, and means and standard deviations

Stressor	Operationalization	Cutoff Percentile Value Used	Mean[a]	Std. Dev
Advanced maternal age	Mother's age at birth in top 10th percentile	38 years	0.116	0.320
Advanced paternal age	Father's age at birth in top 10th percentile	43 years	0.110	0.313
Large sibship size	Number of full siblings (including self) in top 10th percentile	10 siblings	0.164	0.370
Early maternal death	Mother died before ego reached age 18	n/a	0.051	0.221
Early paternal death	Father died before ego reached age 18	n/a	0.077	0.266
High infant mortality rate (IMR) for county of birth	County IMR in top 10th percentile	131.39 per 1,000	0.030	0.172
Late birth order	Birth order in top 10th percentile	6th child	0.197	0.398
Sibling died as infant	At least one sibling died under age 1	n/a	0.257	0.437
Sibling died as child	At least one sibling died between age 1 and 18	n/a	0.162	0.368
Low initial SES	Father's Nam Powers (NP) Score[b] from Usual Occupation on Utah Death Certificate in bottom 10th percentile	NP score of 29	0.103	0.304
Twin/multiple birth	Ego shares same birth year and month with at least one maternal sibling	n/a	0.028	0.165
Out-of-wedlock birth	Parents' marriage year after ego's birth year	n/a	0.004	0.066
Short birth interval	Not firstborn, and number of months since previous sibling's birth in bottom 10th percentile	16 months	0.082	0.274

Notes:

a. Variables are dichotomous, and thus the mean is also the frequency proportion.
b. Nam Powers score converts occupation, a qualitative measure, to a quantitative SES score ranging from 1-100 (Low to high) (*37*).

The sample was then divided into three subsamples: two quarter samples of 134,213 each and one half subsample of 268,424. We plan to utilize the two quarter samples to conduct exploratory factor analysis (EFA) and the half subsample for validation through confirmatory factor analysis (CFA). This approach of EFA followed by CFA in a different sample has been endorsed for scale development (*38*). The best-fitting model following CFA will then be utilized for construction of the ELCS scale. We have already conducted preliminary EFA on the first quarter subsample of 134,213. A Kaiser-Meyer-Olkin test for factorability yielded an overall measure of sampling adequacy (MSA) of 0.68, suggesting that factor analytic methods are appropriate (*38*).

Several key decisions are involved in performing EFA (*38, 39*). One of the most important decisions is the number of underlying factors to retain—the dimensionality of the data (*39, 40*). This is because misspecification of the number of factors tends to severely alter the factor loadings and structure, while other decisions such as factor extraction type or rotation method are more robust across various procedures (*40*).

Parallel Analysis (PA) has been shown to be the most accurate method for determining the correct number of factors, primarily because it accounts for sampling error (*40*) and relaxes the requirement for normally-distributed variables by invoking the Central Limit Theorem (*41, 42*). It is recommended by journal editors as a preferred method (*39, 43*). PA compares the ordered (high to low) eigenvalues from the actual dataset to a null distribution of ordered eigenvalues randomly generated from datasets with the same variables and number of observations. If the factor's actual eigenvalue is greater than the simulated (i.e., the observed greater than what is expected by chance alone), then that factor is retained. For example, if the actual eigenvalue for a given factor is greater than the 95^{th} percentile for the simulated factor, it

can be retained with a confidence level of 1-.05 (where .05 is the chosen alpha) (*41*). For more details beyond this terse summary, the reader is referred to other sources (*40, 41, 44*).

We implemented PA in SAS 9.4 using the *%PARALLEL* macro provided by Kabacoff (*45*), the exceptions being we randomized our variables for only two possible values (0 or 1), and implemented minor aesthetic chart adjustments. We used 5000 iterations and the 95th percentile of the simulated eigenvalues, which represents a conservative estimate (*40*) . Next, in order to determine the general pattern of factor loadings, we performed a principal components analysis (PCA) via PROC FACTOR.

PRELIMINARY FINDINGS

Figure 1 displays the results of our PA in a scree plot, by comparing the actual eigenvalues to the 95th percentiles of the simulated eigenvalues. The values themselves appear in Table 3. Note that four factors have actual eigenvalues higher than the simulated 95th percentiles; therefore, using an alpha=.05, we should retain four factors. The factor pattern resulting from PCA using four factors is shown in Table 4.

Figure 1. Parallel analysis scree plot

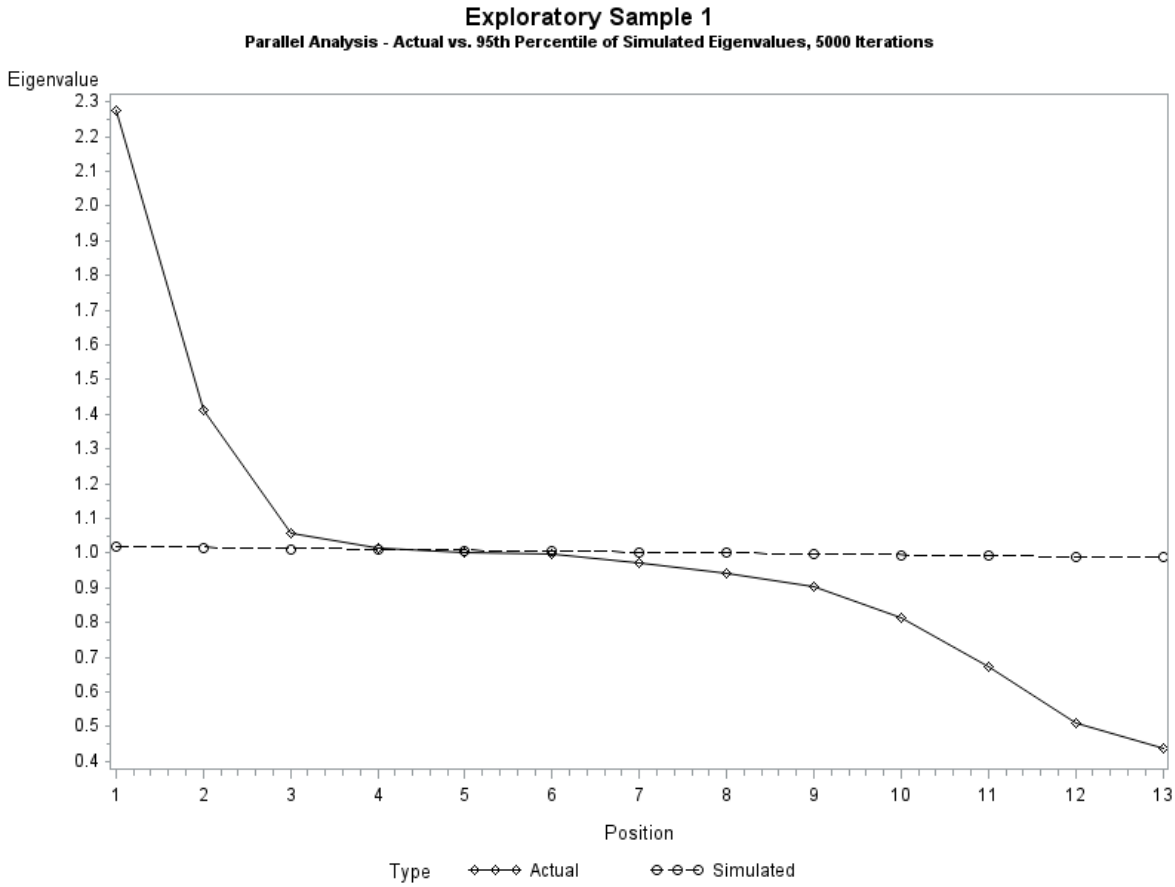

Table 3. Actual vs. 95th percentiles of simulated eigenvalues from parallel analysis with 5000 simulations

Factor Position	Actual Eigenvalue	95th Percentile of Simulated Eigenvalues
F1	2.276	1.019
F2	1.411	1.015
F3	1.056	1.012
F4	1.013	1.009
F5	1.001	1.006
F6	0.998	1.004
F7	0.973	1.002
F8	0.940	1.000
F9	0.902	0.997
F10	0.813	0.995
F11	0.671	0.993
F12	0.509	0.990
F13	0.436	0.987

Note:
Factors where the actual eigenvalues were greater than the 95th percentiles of the simulated are shaded gray.

Table 4. Factor patterns from principal components analysis with four factors; and explained variance

	F1	F2	F3	F4
Factor Patterns				
Advanced maternal age	.83 *	.01	.01	.01
Advanced paternal age	.81 *	.01	.07	.00
Large sibship size	.21	.76 *	-.05	-.05
Early maternal death	.02	.06	.29	.64 *
Early paternal death	.27	-.14	.53 *	.10
High infant mortality rate for county of birth	-.11	.29	-.02	.41 *
Late birth order	.62 *	.48 *	.00	.00
Sibling died as infant	.04	.65 *	.11	.01
Sibling died as child	.08	.58 *	-.03	.12
Low initial SES	-.06	.00	.74 *	-.05
Twin/multiple birth	-.04	.17	-.02	.01
Out-of-wedlock birth	-.03	.02	.12	-.55 *
Short birth interval	-.18	.24	.35	-.36
Variance Explained	1.90	1.76	1.07	1.03
Note: Values greater than .40 are flagged by '*'				

NEXT STEPS

We plan to:

1. Finish the first EFA
    a. Refine the extraction and rotation methods, and delete unnecessary items as appropriate
    b. Interpret factors
2. Complete EFA on the second subsample
3. Using the remaining half sample, perform CFA with both EFA solutions to verify a final best solution, and provide appropriate fit indices
4. Use the final solution to create the ELCS score for each individual
5. Link these scores with longevity and suicide data to model all-cause and suicide mortality hazards with Cox regression models, controlling for potential confounders

REFERENCES

1. I. T. Elo, S. H. Preston, Effects of early-life conditions on adult mortality: a review. *Population index*, 186-212 (1992).
2. K. R. Smith, G. P. Mineau, G. Garibotti, R. Kerber, Effects of childhood and middle-adulthood family conditions on later-life mortality: Evidence from the Utah Population Database, 1850–2002. *Social Science & Medicine* **68**, 1649-1658 (2009); published online Epub5// (http://dx.doi.org/10.1016/j.socscimed.2009.02.010).
3. L. A. Gavrilov, N. S. Gavrilova, Early-Life Programming of Aging and Longevity: The Idea of High Initial Damage Load (the HIDL Hypothesis). *Annals of the New York Academy of Sciences* **1019**, 496-501 (2004)10.1196/annals.1297.091).
4. L. A. Gavrilov, N. S. Gavrilova, When fatherhood should stop? *Science* **277**, 17-21 (1997).
5. J. Mirowsky, Age at First Birth, Health, and Mortality. *Journal of Health and Social Behavior* **46**, 32-50 (2005); published online EpubMarch 1, 2005 (10.1177/002214650504600104).
6. B. Miller, A. Alaraisanen, J. Miettunen, M. R. Jarvelin, H. Koponen, P. Rasanen, M. Isohanni, B. Kirkpatrick, Advanced paternal age, mortality, and suicide in the general population. *The Journal of nervous and mental disease* **198**, 404-411 (2010); published online EpubJun (10.1097/NMD.0b013e3181e07d79).
7. Y.-Y. Chen, D. Gunnell, C.-L. Lu, S.-S. Chang, T.-H. Lu, C.-Y. Li, Perinatal risk factors for suicide in young adults in Taiwan. *International Journal of Epidemiology* **42**, 1381-1389 (2013); published online EpubOctober 1, 2013 (10.1093/ije/dyt129).
8. C. L. Hart, G. Davey Smith, Relation between number of siblings and adult mortality and stroke risk: 25 year follow up of men in the Collaborative study. *Journal of epidemiology and community health* **57**, 385-391 (2003); published online EpubMay 1, 2003 (10.1136/jech.57.5.385).
9. D. Riordan, C. Morris, J. Hattie, C. Stark, Family size and perinatal circumstances, as mental health risk factors in a Scottish birth cohort. *Soc Psychiatry Psychiatr Epidemiol* **47**, 975-983 (2012); published online Epub2012/06/01 (10.1007/s00127-011-0405-5).
10. K. R. Smith, H. A. Hanson, M. C. Norton, M. S. Hollingshaus, G. P. Mineau, Survival of offspring who experience early parental death: Early life conditions and later-life mortality. *Social Science & Medicine*, http://dx.doi.org/10.1016/j.socscimed.2013.11.054).
11. T. Andersson, U. Hogberg, S. Åkerman, Survival of orphans in 19th century Sweden—the importance of remarriages. *Acta Pædiatrica* **85**, 981-985 (1996)10.1111/j.1651-2227.1996.tb14198.x).
12. T. Niederkrotenthaler, B. Floderus, K. Alexanderson, F. Rasmussen, E. Mittendorfer-Rutz, Exposure to parental mortality and markers of morbidity, and the risks of attempted and completed suicide in offspring: an analysis of sensitive life periods. *Journal of epidemiology and community health* **66**, 233-239 (2012); published online EpubMarch 1, 2012 (10.1136/jech.2010.109595).
13. E. Agerbo, M. Nordentoft, P. B. Mortensen, Familial, psychiatric, and socioeconomic risk factors for suicide in young people: nested case-control study. *British Medical Journal* **325**, 74-77 (2002).

14. C. E. Finch, E. M. Crimmins, Inflammatory Exposure and Historical Changes in Human Life-Spans. *Science* **305**, 1736-1739 (2004); published online EpubSeptember 17, 2004 (10.1126/science.1092556).
15. C. E. Finch, E. M. Crimmins, Response to Comment on "Inflammatory Exposure and Historical Changes in Human Life-Spans". *Science* **308**, 1743 (2005); published online EpubJune 17, 2005 (10.1126/science.1109017).
16. D. J. P. Barker, C. Osmond, I. Rodin, C. H. D. Fall, P. D. Winter, *Low weight gain in infancy and suicide in adult life*. (1995), vol. 311, pp. 1203.
17. S. Erhardt, C. K. Lim, K. R. Linderholm, S. Janelidze, D. Lindqvist, M. Samuelsson, K. Lundberg, T. T. Postolache, L. Traskman-Bendz, G. J. Guillemin, L. Brundin, Connecting inflammation with glutamate agonism in suicidality. *Neuropsychopharmacology* **38**, 743-752 (2013); published online Epub04//print (http://www.nature.com/npp/journal/v38/n5/suppinfo/npp2012248s1.html).
18. M. E. Sublette, H. C. Galfalvy, D. Fuchs, M. Lapidus, M. F. Grunebaum, M. A. Oquendo, J. John Mann, T. T. Postolache, Plasma kynurenine levels are elevated in suicide attempters with major depressive disorder. *Brain, Behavior, and Immunity* **25**, 1272-1278 (2011); published online Epub8// (http://dx.doi.org/10.1016/j.bbi.2011.05.002).
19. T. Pace, T. Mletzko, O. Alagbe, D. Musselman, C. Nemeroff, A. Miller, C. Heim, Increased stress-induced inflammatory responses in male patients with major depression and increased early life stress. *American Journal of Psychiatry* **163**, 1630-1633 (2006).
20. L. Salk, W. Sturner, L. Lipsitt, B. Reilly, R. Levat, RELATIONSHIP OF MATERNAL AND PERINATAL CONDITIONS TO EVENTUAL ADOLESCENT SUICIDE. *The Lancet* **325**, 624-627 (1985); published online Epub3/16/ (http://dx.doi.org/10.1016/S0140-6736(85)92156-7).
21. B. Modin, Birth order and mortality: a life-long follow-up of 14,200 boys and girls born in early 20th century Sweden. *Social Science & Medicine* **54**, 1051-1064 (2002); published online Epub4// (http://dx.doi.org/10.1016/S0277-9536(01)00080-6).
22. F. Nault, B. Desjardins, J. Légaré, Effects of Reproductive Behaviour on Infant Mortality of French-Canadians During the Seventeenth and Eighteenth Centuries. *Population Studies* **44**, 273-285 (1990); published online Epub1990/07/01 (10.1080/0032472031000144596).
23. J. H. Bjorngaard, O. Bjerkeset, L. Vatten, I. Janszky, D. Gunnell, P. Romundstad, Maternal age at child birth, birth order, and suicide at a young age: a sibling comparison. *American journal of epidemiology* **177**, 638-644 (2013); published online EpubApr 1 (10.1093/aje/kwt014).
24. M. Rostila, J. Saarela, I. Kawachi, Birth order and suicide in adulthood: evidence from Swedish population data. *American journal of epidemiology* **179**, 1450-1457 (2014); published online EpubJun 15 (10.1093/aje/kwu090).
25. H. M. Gravseth, L. Mehlum, T. Bjerkedal, P. Kristensen, Suicide in young Norwegians in a life course perspective: population-based cohort study. *Journal of epidemiology and community health* **64**, 407-412 (2010); published online EpubMay (10.1136/jech.2008.083485).
26. M. Rostila, J. Saarela, I. Kawachi, The Forgotten Griever: A Nationwide Follow-up Study of Mortality Subsequent to the Death of a Sibling. *American journal of*

epidemiology **176**, 338-346 (2012); published online EpubAugust 15, 2012 (10.1093/aje/kws163).
27. M. Rostila, J. Saarela, I. Kawachi, Suicide following the death of a sibling: a nationwide follow-up study from Sweden. *BMJ Open* **3**, (2013); published online EpubJanuary 1, 2013 (10.1136/bmjopen-2013-002618).
28. B. Galobardes, J. W. Lynch, G. D. Smith, Is the association between childhood socioeconomic circumstances and cause-specific mortality established? Update of a systematic review. *Journal of epidemiology and community health* **62**, 387-390 (2008); published online EpubMay 1, 2008 (10.1136/jech.2007.065508).
29. M. M. Braun, A. Ahlbom, B. Floderus, L. A. Brinton, R. N. Hoover, Effect of twinship on incidence of cancer of the testis, breast, and other sites (Sweden). *Cancer Causes & Control* **6**, 519-524 (1995).
30. J. Baird, C. Osmond, I. Bowes, D. I. W. Phillips, Mortality from birth to adult life: a longitudinal study of twins. *Early Human Development* **53**, 73-79 10.1016/S0378-3782(98)00044-9).
31. C. Tomassini, K. Juel, N. V. Holm, A. Skytthe, K. Christensen, Risk of suicide in twins: 51 year follow up study. *BMJ* **327**, 373-374 (2003); published online Epub2003-08-14 00:00:00 (10.1136/bmj.327.7411.373).
32. M. Voracek, Risk of suicide in twins: Intelligence of twins and singletons may explain risk. *BMJ* **327**, 1168 (2003); published online Epub2003-11-13 00:00:00 (10.1136/bmj.327.7424.1168-a).
33. S. Fors, C. Lennartsson, O. Lundberg, Live long and prosper? Childhood living conditions, marital status, social class in adulthood and mortality during mid-life: A cohort study. *Scandinavian Journal of Public Health* **39**, 179-186 (2011); published online EpubMarch 1, 2011 (10.1177/1403494810395823).
34. J. Davanzo, L. Hale, A. Razzaque, M. Rahman, The effects of pregnancy spacing on infant and child mortality in Matlab, Bangladesh: How they vary by the type of pregnancy outcome that began the interval. *Population Studies* **62**, 131-154 (2008); published online Epub2008/07/01 (10.1080/00324720802022089).
35. D. V. Riordan, C. Morris, J. Hattie, C. Stark, Interbirth spacing and offspring mental health outcomes. *Psychological medicine* **42**, 2511-2521 (2012)doi:10.1017/S0033291712000888).
36. J. E. Wylie, G. P. Mineau, Biomedical databases: protecting privacy and promoting research. *Trends in Biotechnology* **21**, 113-116 (2003); published online Epub3// (http://dx.doi.org/10.1016/S0167-7799(02)00039-2).
37. C. B. Nam, M. G. Powers, *The socioeconomic approach to status measurement* (Cap and Gown Press, Houston, TX, 1983).
38. R. L. Worthington, T. A. Whittaker, Scale Development Research: A Content Analysis and Recommendations for Best Practices. *The Counseling Psychologist* **34**, 806-838 (2006); published online EpubNovember 1, 2006 (10.1177/0011000006288127).
39. T. A. Schmitt, Current Methodological Considerations in Exploratory and Confirmatory Factor Analysis. *Journal of Psychoeducational Assessment* **29**, 304-321 (2011); published online EpubAugust 1, 2011 (10.1177/0734282911406653).
40. J. C. Hayton, D. G. Allen, V. Scarpello, Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis. *Organizational Research Methods* **7**, 191-205 (2004); published online EpubApril 1, 2004 (10.1177/1094428104263675).

41. L. W. Glorfeld, An Improvement on Horn's Parallel Analysis Methodology for Selecting the Correct Number of Factors to Retain. *Educational and Psychological Measurement* **55**, 377-393 (1995); published online EpubJune 1, 1995 (10.1177/0013164495055003002).
42. A. Dinno, Exploring the Sensitivity of Horn's Parallel Analysis to the Distributional Form of Random Data. *Multivariate Behavioral Research* **44**, 362-388 (2009); published online Epub2009/06/22 (10.1080/00273170902938969).
43. R. K. Henson, J. K. Roberts, Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice. *Educational and Psychological Measurement* **66**, 393-416 (2006); published online EpubJune 1, 2006 (10.1177/0013164405282485).
44. J. Horn, A rationale and test for the number of factors in factor analysis. *Psychometrika* **30**, 179-185 (1965); published online Epub1965/06/01 (10.1007/BF02289447).
45. R. I. Kabacoff, Determining the Dimensionality of Data: A SAS® Macro for Parallel Analysis. *SUGI (SAS® Users Group International)*, 1-3 (Paper 90-28).