Taming the Shrewdness of Neural Function: Methodological Challenges in Computational Psychiatry

August 31, 2015

Peter Dayan[1] Raymond J Dolan[2,3] Karl J Friston[2] P Read Montague[2,4]

1. Gatsby Computational Neuroscience Unit, UCL dayan@gatsby.ucl.ac.uk
2. Wellcome Trust Centre for Neuroimaging, UCL
3. Max Planck UCL Centre for Computational Psychiatry and Ageing Research, UCL
4. Human Neuroimaging Laboratory, Virginia Tech Carilion Research Institute

Abstract

Computational psychiatry involves applying a collection of theoretical notions, including data analysis and mathematical and computational modeling, to the problems of psychiatry. It is a nascent field whose central methods are just in the process of being developed. We consider some of the challenges and opportunities for techniques and approaches that are presenting themselves as it starts to take on a more concrete form.

1 Introduction

The field of computational neuroscience [1, 2] has three main facets: (i) data analysis, which provides mostly statistical and machine learning-based techniques for manipulating and understanding the ever-growing wealth of empirical data that it is now possible to collect [3, 4]; (ii) mathematical modeling, which provides for multi-scale treatments of neural phenomena, explaining findings at one level of characterization by (typically quantitative) reduction to mechanisms at lower levels [5]; and (iii) computational modeling, which derives substantial constraints for neural processing from the fact that brains perform information processing functions – i.e., the phenomena play computational roles.

As soon as investigators started to build such mathematical and computational models of normal neural structure and function, the idea that these formal characterizations might illuminate abnormalities such as those apparent in neurology and psychiatry (and indeed vice-versa) was born [6–8]. It was as computational neuroscience started to mature, and, simultaneously, dissatisfaction with the state of psychiatry started to fester, that notions of a more fully-fledged field of computational psychiatry became concrete.

By now, each of the three facets has found some resonance in psychiatry: data analysis, simultaneously reaching a zenith and nadir in psychiatric genetics [9]; mathematical

modeling, for instance evident in the analysis of altered network dynamics associated with imbalances between excitation and inhibition [10]; and computational modeling, in the extensive investigations of disordered decision-making [11].

These successes in turn have led to a number of enthusiastic reviews (or somewhat more accurately, previews) of the field [12–17], including some by various of us. However, a body of clear and compelling methods is a key preliminary to the sort of new understanding and nosology (i.e., systematic classification) of psychiatric conditions that are popular interim goals in the field, let alone to the potential therapeutic advances that even the brave are as yet far from offering.

In this review, we consider some of the existing and desirable methodological steps for the field. Most methods are not unique to psychiatry – they just need careful application. However, some, for instance to do with individual differences, are of more immediate significance in psychiatry than in some other neuroscience disciplines. Given limited space, and the modeling focus of the panoply of previews, we mainly focus on data analysis, touching only briefly on relevant aspects of the two forms of modelling. Of the many areas in which methods of data analysis are playing, or could play, a crucial role in computational psychiatry, two of very general importance concern (a) dimensionality reduction and more general ways of finding statistical structure in very high dimensional data; and (b) a specially noteworthy case of dimensionality reduction, namely ways of characterizing differences within and between populations, at both single points in time, and longitudinally.

2 Taming complexity through low dimensional structure

The bewildering complexity of the anatomy and physiology of the nervous system, together with those of its genetic and environmental determinants, require substantial taming in order for it to be possible to make progress in understanding what can go right and wrong. Taming is typically understood in terms of finding low dimensional structure that quantitatively and/or qualitatively characterizes central aspects of the full problem, or at least provides a path to a form of sequential expansion.

As an illustrative example, one of the most vibrant areas of experimental and theoretical research concerns mis-wiring - abnormalities and disease as a form of functional dysconnection or synaptopathy [18]. Wiring can be wrongly or additionally routed (as suggested, for instance, in synaesthesia [19]), or over-exuberant or under-pruned, at least some times over the course of development; there could be more fine-scale problems, such as the make-up of subunits of membrane-embedded channels (as in channelopathies; [20]), or indeed the nature of synaptic plasticity, which adjusts these characteristics typically over the course of the interaction between the individual and their environment [21, 22].

The first step in any of these directions involves being able to assess normal and abnormal states of wiring. Network analysis methods - the qualitative understanding of patterns of connectivity (small-worldness; hubs and the like; [23]) provide just such a characterization - structure in the gargantuan space of connectivity matrices. Ideas for the implications of such qualitative structures for the flow of information in the brain remain in demand [24, 25].

More prosaically, even the first stages of any analysis of structure - the determination of what is connected to what and by what means, poses a monumental challenge - methods that facilitate or augment manual segmentation of images from electron microscopy in order to determine the nature of the connections [26] are of obvious note.

What goes for anatomy also applies to physiology. Again, just as an example, there has been much work considering the low-dimensional structure in the dynamics of the activity of large populations of neurons – enabled by recent advances in methods for large-scale simultaneous electrical or optical recording. There are various reasons to think that such structure will exist - for instance, it has been noted (Ganguli, personal communication) that the dimensionality of the input or output that are encoded is often very modest compared with the huge number of neurons. Qualitative structures such as surprisingly sluggish low-dimensional attractor dynamics [27] have been extracted using advanced statistical methods from multi-unit recordings; these turn out to have implications for behaviourally measurable quantities such as reaction times and various forms of variability. It is, however, early days for our understanding of the nature and functional role of such structures across different spatial and temporal scales.

Other suggestions, such as chaotic itineracy [28, 29] – that such low dimensional state rove substantially over a whole domain – have been tied to abnormalities. More generally, so-called 'dynamical diseases' [30–32] are supposed to arise if the state evolves in an unusual manner, visiting potentially incorrect regions of state space in an incorrectly controlled way. Methods such as dynamic causal modeling (DCM) based on effective or functional connectivity are also starting to prove their mettle as ways of divining various aspects of abnormalities [33].

Mathematical modeling would ideally provide the link between these anatomical and physiological cases, answering how patterns of wiring, together with the characteristics of the neural elements that are thereby coupled, lead to the dynamics of activity that are observed [34, 35]. More generally, mathematical modeling provides a form of multiscale analysis, associated with the huge range of temporal and spatial scales [2] that are relevant for the brain. This is of particular value in trying to understand cause and effect - something that is critical to get at the heart of the problems associated with disease.

There has perhaps been rather less computational modeling associated with these qualitative characteristics. One main exception concerns attractor models, which have been implicated in a host of computational operations. The effect of abnormal (e.g., dopaminergic) neuromodulation, for instance in schizophrenia, in the dynamics of such networks has been implicated in computationally characterised aspects of the disorder such as ab-

normal fixedness and flightiness [6]. The idea is that the gain of neurons (i.e., the slope of the input-output relationship) is modulated by neuromodulators. The effect of this is either to over- or under-stabilize points of attraction, which themselves represent states of cognitive importance such as goals or short-term memories. Aspects of oscillations have also been awarded computational roles, albeit as yet with rather nascent links to psychiatry [36].

3 Individual differences

Consider any way whatsoever of assessing genes, anatomy, physiology, or indeed behaviour across one or more populations of individuals. It is an obvious truism that the study of dysfunction must begin with a characterization of the way that these facets vary within and between these groups, and indeed the longitudinal reliability of the instruments used for this characterization [37]. For a start, it is impossible to define the abnormal without reference to the normal. More subtle, though, are the forms of structure prevalent in the populations. This has implications for such things as categorical or discrete versus spectrum conditions [38–40], and also temporal characteristics evident in the familiar distinction between traits and states [41]. However, it is also of note in periodic diseases in psychiatry such as bipolar disorder and others [31], and secular changes as in development, ageing and indeed dementing disorders, for which there are sophisticated statistical treatments [42–47], for instance involving forms of structural equation modeling.

Some such characterizations are commonplace – for instance, principal components and factor analysis are in ubiquitous use. These can be seen as assuming a particular sort of Gaussian characterization of the variables concerned (e.g., questionairre measures) [48, 49], and finding a typically restricted number of axes associated with their covariance matrices. Each axis defines a spectrum, realizing continuous dimensions of variability.

However, continuous spectra are not the only possibility. One could equally perform clustering, as in statistical mixture models [50], which can quite naturally lead to the notion of discrete disorders. They can also be naturally combined with dimensional models as in mixtures of factor analysers [51, 52], a model that applies if the dimensional or spectral structure within different discrete clusters is different.

There are many methods for discovering, validating and testing such so-called latent variable models [53]. These get their name from the fact that the aspect of the structure that underlies each example, for instance the cluster whence it hails, is not a direct part of the input, but is rather latent or hidden and has to be discovered. These are often seen as random effects models – since individuals, individual examples, or, more richly, individual (neural) mechanisms or systems are seen as typically independent samples drawn from an underlying population distribution.

[33] provides an inspiring example in the case of schizophrenia, involving a reanalysis of fMRI data from patients and controls performing a simple visual working memory task. These authors performed clustering, using a statistically sophisticated rendition of the data, and found not only that those with a disorder naturally separated from those without, but also that the patients could be separated further into three subgroups in an unsupervised manner, using only the pattern of their neural responses.

One popular approach for fitting such models is maximum likelihood density estimation, for instance using the expectation maximization algorithm. In cases in which the actual input is itself a noisy or partial reflection of the underlying parameters (something that happens routinely when the underlying data being fit are *parameters* of a behavioural model, such as the learning rate or the sensitivity to reward, and the input are observed choices), it may be necessary to build more layers of latent structure, and to use approximations to perform the fitting. Of particular importance in these cases is model comparison – assessing which model fits the data better can provide a (typically, and importantly, incomplete) statistical justification for claims about the structure of a disease or a class of diseases. More complex models with more parameters can typically fit data more accurately; thus proper comparison requires complexity to be correctly penalised. There are various ways of doing this - notable examples are using hold-out data, or other forms of cross-validation [54], and approximate Bayesian methods such as the Bayesian Information Criterion or the Akaike Information Criterion [55].

An increasingly popular alternative to maximum likelihood fitting is to employ a more fully-fledged Bayesian approach [56–59]. This can make fewer approximations, but at the expense of greater computational cost, for instance accrued by Markov chain Monte-Carlo sampling [59]. One particular advantage of the Bayesian methods is that they more readily afford the possibility of what are known as non-parametric models, i.e., avoiding *a priori* restrictions on such things as the number of clusters or factors. They also automatically penalize complexity, via a form of Occam's razor – although there are various theoretical concerns with model comparison, as Bayesians prefer to average over models rather than select between them. Averaging is based on the marginal likelihood or model evidence, which, unfortunately, not only is often very hard to compute, but can also depend on a set of assumptions about prior distributions whose justification may not always be completely transparent.

In view of these various uncertainties, investigators often use multiple methods to assess population differences [33, 60]. One method is to compare mixture models fit in both supervised and unsupervised ways - i.e., with and without knowledge of the putative population structure (control or diseased). Another is to calculate and compare summary statistics of the exact or approximate posterior distributions over the implied parameters or characteristics.

It is also important to study the structure of outliers – i.e., extremes relative to any of these distributions [61]. Outliers pose many statistical problems (not the least because of the conventional focus on eliminating them as noise rather than studying them as sig-

nal; [62]). Investigating them is significantly dependent on getting access to very large populations.

Unlike more mature fields such as development and ageing [63–66], computational psychiatry has yet to come firmly to grips with longitudinal aspects of the characteristics of its populations – capturing the various sorts of changes that can occur across time, and indeed finding signs pre-morbidly that can be of clinical benefit. This should not only include evidence related to shorter-term states and longer-term traits, but also more basic questions such as test-retest reliability (which could be importantly affected by recall of the reinforcement contingencies, or meta-contingencies such as the speed of change in the task). There is a dearth of work on more sophisticated aspects of directed change over time that require more comprehensive longitudinal models [42–47, 67].

4 Discussion

Many of the initial efforts in computational psychiatry concerned ideas to do with decision-making. This is partly because this is a key aspect of information processing that is disturbed in psychiatric conditions. It is also because decision-making is an area in which there are powerful models of normative function that link computations with psychological and neural findings, and also environmental influences on information processing (in the shape of priors, with consequences for biases, generalization and more; [11]).

However, the other aspects of computational psychiatry are of at least equal importance. Mathematical modeling is necessary to provide multiscale analyses that can tie malfunctioning or mis-wired elements to their dynamical consequences. Data analysis, on which we focused here, is critical to provide compact, and thus revealing, analyses of the otherwise overwhelming complexity of the brain. It is also essential to provide an analysis of the structure within and between populations, to help delimit abnormalities.

There remain a wealth of areas requiring further work and analysis. Prime amongst these is to import data analytical ideas about systematic change into the field to capture secular and oscillatory evolution, along with the bio- and behavioural markers which will have the appropriate discriminative and generative capacities.

Acknowledgements

PD is funded by the Gatsby Charitable Foundation; RJD by a Wellcome Trust Senior Investigator Award (091593/Z/10/Z); KJF by a Wellcome Trust Principal Research Fellowship (Ref: 088130/Z/09/Z); PRM by National Institutes of Health grants RC4 AG039067, R01 DA11723, MH 085496, The Kane Family Foundation, and a Principal Research Fellowship from The Wellcome Trust.

References

[1] Schwartz EL: *Computational neuroscience*. Mit Press; 1993.

[2] Sejnowski TJ, Koch C, Churchland PS: **Computational neuroscience**. *Science* 1988, **241**:1299–1306.

[3] Kass RE, Eden UT, Brown EN: *Analysis of neural data*. Springer; 2014.

Tremendous didactic resource for methods of neural data analysis.

[4] Sejnowski TJ, Churchland PS, Movshon JA: **Putting big data to good use in neuroscience.** *Nat Neurosci* 2014, **17**:1440–1441.

A call to arms for 'big data' in neuroscience – both in terms of collection and analysis.

[5] Ermentrout GB, Terman DH: *Mathematical foundations of neuroscience*, volume 35. Springer Science & Business Media; 2010.

[6] Cohen JD, Servan-Schreiber D: **Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia.** *Psychol Rev* 1992, **99**:45–77.

[7] Hoffman RE: **Computer simulations of neural information processing and the schizophrenia-mania dichotomy.** *Arch Gen Psychiatry* 1987, **44**:178–188.

[8] Braver TS, Barch DM, Cohen JD: **Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function.** *Biol Psychiatry* 1999, **46**:312–328.

[9] Flint J, Goodwin G: **Psychiatric genetics: a genetic basis for health?** *Curr Biol* 1999, **9**:R326–R328.

[10] Eichler SA, Meier JC: **E-i balance and human diseases - from molecules to networking.** *Front Mol Neurosci* 2008, **1**:2.

[11] Huys Q, Guitart-Masip M, Dolan R, Dayan P: **Decision-theoretic psychiatry**. *Clinical Psychological Science* 2015, **2167702614562040**.

[12] Tretter F, Gebicke-Haerter PJ: **Systems biology in psychiatric research: from complex data sets over wiring diagrams to computer simulations.** *Methods Mol Biol* 2012, **829**:567–592.

[13] Maia TV, Frank MJ: **From reinforcement learning models to psychiatric and neurological disorders.** *Nat Neurosci* 2011, **14**:154–162.

[14] Montague PR, Dolan RJ, Friston KJ, Dayan P: **Computational psychiatry.** *Trends Cogn Sci* 2012, **16**:72–80.

[15] Huys QJM, Moutoussis M, Williams J: **Are computational models of any use to psychiatry?** *Neural Netw* 2011, **24**:544–551.

[16] Wang XJ, Krystal JH: **Computational psychiatry.** *Neuron* 2014, **84**:638–654.

[17] Stephan KE, Mathys C: **Computational approaches to psychiatry.** *Curr Opin Neurobiol* 2014, **25**:85–92.

[18] Di Martino A, Fair DA, Kelly C, Satterthwaite TD, Castellanos FX, Thomason ME, Craddock RC, Luna B, Leventhal BL, Zuo XN, Milham MP: **Unraveling the miswired connectome: a developmental perspective.** *Neuron* 2014, **83**:1335–1353.

Reviews data and ideas on mis-wiring over the course of development, both structural (revealed, for instance, by diffusiuon-based imaging) and functional (with a focus on resting-state functional imaging.

[19] Bargary G, Mitchell KJ: **Synaesthesia and cortical connectivity.** *Trends Neurosci* 2008, **31**:335–342.

[20] Gargus JJ: **Ion channel functional candidate genes in multigenic neuropsychiatric disease.** *Biol Psychiatry* 2006, **60**:177–185.

[21] Meyer-Lindenberg A, Tost H: **Neuroimaging and plasticity in schizophrenia.** *Restor Neurol Neurosci* 2014, **32**:119–127.

[22] Wondolowski J, Dickman D: **Emerging links between homeostatic synaptic plasticity and neurological disease.** *Front Cell Neurosci* 2013, **7**:223.

[23] Newman M, Barabasi AL, Watts DJ: *The structure and dynamics of networks*. Princeton University Press; 2006.

[24] Bullmore E, Sporns O: **Complex brain networks: graph theoretical analysis of structural and functional systems**. *Nature Reviews Neuroscience* 2009, **10**:186–198.

[25] Sporns O: *Networks of the Brain*. MIT press; 2011.

[26] Jain V, Seung HS, Turaga SC: **Machines that learn to segment images: a crucial technology for connectomics**. *Current opinion in neurobiology* 2010, **20**:653–666.

[27] Shenoy KV, Sahani M, Churchland MM: **Cortical control of arm movements: a dynamical systems perspective.** *Annu Rev Neurosci* 2013, **36**:337–359.

Considers the structure of the activity of populations of neurons in the revealing terms of dynamical systems, with links to control theoretic ideas.

[28] Tsuda I: **Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems**. *Behavioral and Brain Sciences* 2001, **24**:793–810.

[29] Kay LM: **A challenge to chaotic itinerancy from brain dynamics.** *Chaos* 2003, **13**:1057–1066.

[30] Globus GG, Arpaia JP: **Psychiatry and the new dynamics.** *Biol Psychiatry* 1994, **35**:352–364.

[31] Mackey MC, Milton JG: **Dynamical diseases.** *Ann N Y Acad Sci* 1987, **504**:16–32.

[32] Tretter F, Gebicke-Haerter PJ, an der Heiden U, Rujescu D, Mewes HW, Turck CW: **Affective disorders as complex dynamic diseases–a perspective from systems biology.** *Pharmacopsychiatry* 2011, **44 Suppl 1**:S2–S8.

[33] Brodersen KH, Deserno L, Schlagenhauf F, Lin Z, Penny WD, Buhmann JM, Stephan KE: **Dissecting psychiatric spectrum disorders by generative embedding.** *Neuroimage Clin* 2014, **4**:98–111.

Generates features for supervised and unsupervised classification and analysis of patients with schizophrenia through dynamic causal modeling of regions involved in a working-memory task. This is an important proof of principle of a number of methods and ideas for both the 'taming complexity' and 'individual differences' parts of this review.

[34] Strogatz SH: *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering*. Westview press; 2014.

[35] Woolrich MW, Stephan KE: **Biophysical network models and the human connectome.** *Neuroimage* 2013, **80**:330–338.

[36] Tretter F, Pogarell O, Rujescu D, Meisenzahl E, Mewes HW: **Systems biology of oscillatory processes in sleep and mental disorders.** *Pharmacopsychiatry* 2013, **46 Suppl 1**:S1.

[37] Heise DR: **Separating reliability and stability in test-retest correlation**. *American sociological review* 1969, 93–101.

[38] Widiger TA, Samuel DB: **Diagnostic categories or dimensions? a question for the diagnostic and statistical manual of mental disorders–fifth edition.** *J Abnorm Psychol* 2005, **114**:494–504.

[39] Markon KE, Krueger RF: **Information-theoretic latent distribution modeling: distinguishing discrete and continuous latent variable models.** *Psychol Methods* 2006, **11**:228–243.

[40] Goldberg D: **Plato versus aristotle: categorical and dimensional models for common mental disorders**. *Comprehensive psychiatry* 2000, **41**:8–13.

[41] Steyer R, Ferring D, Schmitt MJ: **States and traits in psychological assessment.** *European Journal of Psychological Assessment* 1992, 79–98.

[42] Collins LM: **Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model**. *Annu Rev Psychol* 2006, **57**:505–528.

[43] McArdle JJ: **Latent variable modeling of differences and changes with longitudinal data**. *Annual review of psychology* 2009, **60**:577–605.

[44] Proust-Lima C, Amieva H, Jacqmin-Gadda H: **Analysis of multivariate mixed longitudinal data: a flexible latent process approach.** *Br J Math Stat Psychol* 2013, **66**:470–487.

[45] Fitzmaurice G, Davidian M, Verbeke G, Molenberghs G: *Longitudinal data analysis*. CRC Press; 2008.

[46] Bollen KA, Curran PJ: *Latent curve models: A structural equation perspective*, volume 467. John Wiley & Sons; 2006.

[47] Verbeke G, Fieuws S, Molenberghs G, Davidian M: **The analysis of multivariate longitudinal data: A review**. *Statistical methods in medical research* 2014, **23**:42–59.

[48] Floyd FJ, Widaman KF: **Factor analysis in the development and refinement of clinical assessment instruments.** *Psychological assessment* 1995, **7**:286.

[49] Fabrigar LR, Wegener DT, MacCallum RC, Strahan EJ: **Evaluating the use of exploratory factor analysis in psychological research.** *Psychological methods* 1999, **4**:272.

[50] McLachlan G, Peel D: *Finite mixture models*. John Wiley & Sons; 2004.

[51] Gibbons RD, Dorus E, Ostrow DG, Pandey GN, Davis JM, Levy DL: **Mixture distributions in psychiatric research.** *Biol Psychiatry* 1984, **19**:935–961.

[52] Lubke GH, Muthén B: **Investigating population heterogeneity with factor mixture models.** *Psychological methods* 2005, **10**:21.

[53] Skrondal A, Rabe-Hesketh S: *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Crc Press; 2004.

[54] Kohavi R, et al.: **A study of cross-validation and bootstrap for accuracy estimation and model selection**. In *IJCAI*. 1995, volume 14, 1137–1145.

[55] Penny W: **Comparing dynamic causal models using AIC, BIC and free energy.** *Neuroimage* 2012, **59**:319–330.

[56] Kruschke JK: **Bayesian data analysis**. *Wiley Interdisciplinary Reviews: Cognitive Science* 2010, **1**:658–676.

[57] Lee MD, Wagenmakers EJ: *Bayesian cognitive modeling: A practical course*. Cambridge University Press; 2014.

[58] Daw ND: **Trial-by-trial data analysis using computational models**. In *Decision making, affect, and learning: Attention and performance XXIII*, Oxford University Press; 2011, volume 23, 3–38. 2011.

> Impressively instructive chapter describing the details of building amodels for tasks involving trial-by-trial adaptation, and fitting these models to behavioural data using Bayesian methods.

[59] Stan Development Team: *Stan Modeling Language Users Guide and Reference Manual,*. ; 2014.

[60] Moutoussis M, Bentall RP, El-Deredy W, Dayan P: **Bayesian modelling of jumping-to-conclusions bias in delusional patients.** *Cogn Neuropsychiatry* 2011, **16**:422–447.

[61] Mourão-Miranda J, Hardoon DR, Hahn T, Marquand AF, Williams SCR, Shawe-Taylor J, Brammer M: **Patient classification as an outlier detection problem: an application of the one-class support vector machine.** *Neuroimage* 2011, **58**:793–804.

[62] Huber PJ: *Robust statistics*. Springer; 2011.

[63] Hertzog C, Nesselroade JR: **Assessing psychological change in adulthood: an overview of methodological issues.** *Psychology and aging* 2003, **18**:639.

[64] McArdle JJ, Grimm KJ, Hamagami F, Bowles RP, Meredith W: **Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurement.** *Psychol Methods* 2009, **14**:126–149.

[65] Nesselroade JR: **The warp and the woof of the developmental fabric**. In *Visions of aesthetics, the environment, and development: The legacy of Joachim F. Wohlwill*, Edited by Downs RM, Liben LS, Palermo DS, Erlbaum Hillsdale, NJ; 1991, 213–240. 1991.

[66] Lövdén M, Bäckman L, Lindenberger U, Schaefer S, Schmiedek F: **A theoretical framework for the study of adult cognitive plasticity.** *Psychological bulletin* 2010, **136**:659.

[67] Song C, Kuo L, Derby CA, Lipton RB, Hall CB: **Multi-stage transitional models with random effects and their application to the einstein aging study.** *Biom J* 2011, **53**:938–955.