HHS Public Access
Author manuscript
Health Psychol Rev. Author manuscript; available in PMC 2018 September 01.

Published in final edited form as:
Health Psychol Rev. 2017 September ; 11(3): 292–294. doi:10.1080/17437199.2017.1347514.

Envisioning a Future for Precision Health Psychology: Innovative Applied Statistical Approaches to N-of-1 Studies

Karina W Davidson, PhD, MASc[1,2] and **Ying Kuen Cheung, PhD**[3]

[1]Columbia University Irving Medical Center New York, NY

[2]NewYork Presbyterian Hospital, New York, NY

[3]Columbia University Mailman School of Public Health, New York, NY

Health psychology clinicians and patients routinely engage in clinical encounters where they attempt to determine the best behavior change approach for that particular patient. As is recognized frequently by both participants, this encounter is often unsuccessful. As every evidence-based clinician knows, the provision of group randomized controlled trial data and observational epidemiological study results, while crucial, does not necessarily provide relevant information for how the patient in the office right now should be treated, and how they will respond. The evidence we provide for most health psychology problems is derived from between-person designs, or designs that presume that the set of predictors identified for a group apply to the majority (if not all) persons. We further use between-subject designs to test behavior change techniques, and correctly assume that by comparing (with randomization) one groups' response to a behavior change technique to another groups' response to an attention control condition, or to an alternative treatment, we will know how the 'average' person will respond to the behavior change intervention. Even when a treatment is found to be beneficial, on average, when compared to the effect of a control condition, there will be some individual treatment participants who did not benefit, or who even experienced harm. Likewise, there will be some individual control participants who experienced benefit. Only if the distributions of response to the treatment and control conditions are non-overlapping, can one presume that most individual participants will benefit from the treatment. Non-overlapping response distributions is a rare phenomenon in health psychology.

This is in contrast to within-person or N-of-1 designs, in which the set of significant predictors is identified for that one person, the risk/variance associated with that set of predictors is calculated for that one person, and the response to treatment(s) is calculated for that one person. A N-of-1 design can also be used to determine if a personalized model is needed for each person, or if a few models can be used for discrete subgroups of persons, or if only one model is all that is needed for the entire population, thus addressing many

Authors Contact Information: Karina W. Davidson, PhD, 622 W 168th Street, PH9-314, New York, NY 10032, T: 212-342-4493, F: 212-342-3423, Kd2124@cumc.columbia.edu & Ying Kuen Cheung, PhD., PH/Bio 722 W 168 St, R630, New York, NY 10032, T: 212-305-3332, yc632@cumc.columbia.edu.

Financial Disclosure: No relevant conflicts of interest. Karina W Davidson is a member of the United States Preventive Services Task Force (USPSTF). This article does not necessarily represent the views and policies of the USPSTF.

different health psychology problems where patient heterogeneity is ubiquitous. This approach then can contribute to the creation of a field of precision health psychology. Precision treatment targets may operate differently from person to person, in terms of the set of targets that operate for a person, the time lag that operates for that person, and the level of the behavior change treatment (dose-response) needed for each target. Therefore, a main methodological issue for N-of-1 designs is incorporating these factors in the data analytics.

In the current article (cite RHPR-2017-0020), the authors rightfully tackle one of the first barriers we face in making N-of-1 studies available for health psychology clinicians, patients, theoreticians, and the public. The authors suggest using dynamic modeling framework to account for serial dependencies, time-varying covariates, time lag, and periodicity from data from N-of-1 studies. Since the framework is regression model-based, it has the flexibility to address many analytical issues such as error structure variations, different outcome types, and dose-response. To illustrate the use of dynamic regression, the authors examine an individual's data on exercise bouts obtained from a fitness tracker, and characterize daily activity by a binary outcome (e.g., the presence of at least a 10-minute of activity in one day). The authors then expand the model by including exogenous variables (e.g., day of the week), endogenous variables (e.g., retirement, partner's influence, length of sleep), time lag, time trend, and periodicity as predictors for the presence of exercise bouts.

An advantage of a regression model-based framework is the ease in interpretation of the analysis results *quantitatively*. The authors noted that existing analytical approaches for behavioral and psychological N-of-1 studies preferred visual examination of the raw data, with the presumption that useful patterns and variability will be recognizable to the naked eye. This is consistent with the report by Shaffer and others (Shaffer, Falzon, Cheung, & Davidson, 2015) that few N-of-1 studies provided a clear statistical analysis plan. Specifically, the article reports the main findings for the one participant in a multivariable analysis; see Table 2 in the article. The analysis identified an interesting two-day time lag in the effect of sleeping hours on exercise bouts, whereas the pre-specified conjecture that retirement would have an impact was found statistically nonsignificant. While time (e.g., day of the week, period of the day) is a critical element in all N-of-1 studies, the authors found a large and highly statistically significant reduction in physical activity during the evening period—for this individual. Such individualized inference is perhaps the most powerful aspect of dynamic modeling, and also the most misunderstood feature of N-of-1 studies. The fact that the inference is applicable only to an individual is often criticized as lack of generalizability. When interpreting results from N-of-1 studies, it is important to maintain the right perspective that the *main* goal of N-of-1 studies (and more generally patient-oriented implementation research) is not to produce generalizable knowledge, but rather to improve the care of individuals in their current clinical context (Cheung & Duan, 2014)

As pointed out by the authors, many statistical methods have been proposed to deal with data that are collected longitudinally and dynamically as in an N-of-1 study. The approaches include time series formulation, simulation method, and adaptive treatment regimes. (Cheung et al., 2015; Mills, 1991; Nash & Borckardt, 2011) The proposed dynamic regression is an attempt to put forward a simple data analytic solution that allows those interested in obtaining N-of-1 observational or interventional data the opportunity to more appropriately analyze it. Indeed, the model also has the flexibility to extend to accommodate non-linear time trend and different outcome types, and to be integrated with the above-mentioned methods should the occasion require additional sophistication in the analysis.

The authors end with a discussion on sample size determination and appropriately the needs for further research. This is fundamentally a question about controlling the rate of erroneous conclusions from N-of-1 studies, in which sample size refers to the number of data points obtained from an individual (rather than the number of individuals). The conventional evaluation framework aims to control type I and type II errors. This approach has several limitations with a complex variance-covariance structure (e.g., autocorrelation) and many predictors. Should one be conservative, with this amount of statistical testing, and apply strict family-wise correction (e.g. Bonferroni's correction), or exploratory, and allow no correction at all? Can one form an *a priori* hypothesis that makes conventional sense, when one has decided that specific variable may be uniquely related to the important outcome? Alternative evaluation frameworks may be considered in this regard. First, sample size may be determined so as to maximize cost-effectiveness (Cheung and Duan, 2014) by asking the question whether the benefit (increase in precision) of more data outweighs the cost (burden on the individual). Second, N-of-1 decision making can be formulated as a classification problem, (Breiman, 2001) and simulation may be used to inform sample size. This remains a statistical quandary that is not remediated nor aggravated by the use of dynamic regression. Indeed, this will remain so unless we can stay true to the paradigm shift in thinking about the role of N-of-1 studies in the world of health behavior—to generalize to the one person's behavior, and not to all persons' behavior.

N-of-1 studies have the enormous potential to give patients the evidence they need most to make decisions about their health behavior, and provide results that are based on a specific patient's own conditions, symptoms, and diseases; that consider patient's response to context, therapies, and interventions; and that use patient's unique important and chosen outcomes. In this way, N-of-1 studies are the foundational design for a truly patient-centered approach for achieving precision health psychology. Despite the promise, there has been little uptake of this N-of-1 methodology, except for isolated instances where many barriers were encountered. (Kravitz et al., 2008) Statistical approaches, as well as many other issues, must be addressed before this design can truly be user-friendly. The authors of this article bring us one step closer by adding a tool to the arsenal of N-of-1 data analytics.

Acknowledgments

Funding/Support: This work was supported by grants U24AG052175 and R01MH109496 from the National Institutes of Health, Bethesda, MD.

References

Breiman L. Random forests. Machine learning. 2001; 45(1):5–32.

Cheung YK, Duan N. Design of implementation studies for quality improvement programs: an effectiveness-cost-effectiveness framework. American Journal of Public Health. 2014; 104(1):e23–30. DOI: 10.2105/ajph.2013.301579

Cheung YK, Chakraborty B, Davidson KW. Sequential multiple assignment randomized trial (SMART) with adaptive randomization for quality improvement in depression treatment program. Biometrics. 2015; 71:450–459. [PubMed: 25354029]

Hagger, et al. … Dynamic modeling of N-of-1 data: powerful and flexible data analytics applied to individualized studies. Health Psychology Review. (RHPR-2017-0020).

Kravitz RL, Duan N, Niedzinski EJ, Hay MC, Subramanian SK, Weisner TS. What ever happened to N-of-1 trials? Insiders' perspectives and a look to the future. Milbank Quarterly. 2008; 86(4):533–555. DOI: 10.1111/j.1468-0009.2008.00533.x [PubMed: 19120979]

Mills, TC. Time series techniques for economists. Cambridge University Press; 1991.

Nash MR, Borckardt JJ. How to conduct and statistically analyze case-based time series studies, one patient at a time. Journal of Experimental Psychopathology. 2011; 2(2):139–169.

Shaffer JA, Falzon L, Cheung K, Davidson KW. N-of-1 randomized trials for psychological and health behavior outcomes: a systematic review protocol. Systematic Reviews. 2015; 4(1):87. [PubMed: 26081256]