Child Abuse & Neglect 64 (2017) 13–18

Contents lists available at ScienceDirect

Child Abuse & Neglect

Research article

Assessing child maltreatment prevention via administrative data systems: A case example of reproducibility☆

Ronald J. Prinz

Parenting & Family Research Center, University of South Carolina, 1233 Washington Street (2nd Floor), Columbia, SC 29208, United States

ARTICLE INFO

Article history:
Received 29 July 2016
Received in revised form
30 November 2016
Accepted 11 December 2016

Keywords:
Administrative records
Child abuse and neglect
Child maltreatment
Parenting intervention
Place randomized trial
Prevention
Reproducibility

ABSTRACT

Critical issues about scientific reproducibility have been raised about biomedical research, including the reliability of data and analyses within a given study. The case example in this article examined a reproducibility issue pertaining to the use of administrative data systems for evaluation of child maltreatment (CM) prevention, making use of a prevention study conducted over a decade ago that provided a unique opportunity. The place-randomization study, which randomized counties to condition, found that community-wide implementation of a parenting and family support intervention produced positive impact on county-wide rates for substantiated CM cases and out-of-home placements, documented through a state information system. The key consideration is whether and to what extent the administrative record data re-examined retroactively a decade later for the original study's time period would yield comparable results to those based on data acquired at the time of the study. The results indicated that despite small changes over time, the same data patterns and statistical effects were reproducible for the two archival outcome variables. For substantiated CM, the reproduced analyses reflected higher effect sizes and a clear pattern of reduction as a function of intervention. For out-of-home placements, effect sizes were quite comparable to the original ones, reflecting preventive impact. Overall, this case study illustrated the verifiability of data reproducibility in the context of a population outcome evaluation, which underscores the importance of reliable population-prevalence measurement as an essential part of a comprehensive public health strategy aimed at the prevention of CM.

1. Reproducibility

In recent years, many areas of science have raised critical issues about reproducibility in research (Collins & Tabak, 2014; Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016). Reproducibility refers to a complex set of issues pertaining to the conduct and interpretation of research. Across empirical, computational, and statistical realms, reproducibility can mean repeatability, robustness, reliability, and generalizability (Gorski, 2016). The case example in this article addresses a reproducibility issue pertaining to the reliability of data and analyses based on administrative records in the child welfare system.

☆ The research on which this article was based was supported by Grant #U17/CCU422317 from the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention (CDC). The findings and conclusions in this article are those of the author and do not necessarily represent the views of CDC. The author has served as a consultant to Triple P International, which is a joint-venture entity of the University of Queensland responsible for the dissemination of the Triple P system of interventions.

E-mail address: prinz@mailbox.sc.edu

http://dx.doi.org/10.1016/j.chiabu.2016.12.005
0145-2134/

The specific case, which draws on a previously conducted outcome study focused on prevention of child maltreatment, arose because the administrative data system on which the study was based had gone through subtle changes over time.

2. Administrative data systems in child welfare

Data systems in child welfare are of unquestionable value for several purposes such as documenting problems and needs, understanding interrelationships among indicators and geographical variables, assessing preventive intervention outcomes, and gauging effects of policy shifts. One such function pertains to the use of archival records to evaluate impact of child maltreatment (CM) prevention. A central issue underlying this case example is whether state information systems are sufficiently reliable for this important purpose.

By way of background, the utilization of administrative records in the evaluation of interventions is a critical and growing aspiration in many public policy areas. Recent emphasis on evidence-based policy has produced strong advocacy for cost efficient administrative-record evaluation strategies to be applied for example to intervention and prevention in education, juvenile justice, mental health, poverty reduction, and of course child welfare (Baron, 2012; Feeney, Bauman, Chabrier, Mehra, & Woodford, 2015; Haskins & Margolis, 2015).

The ultimate goal of prevention generally is prevalence reduction, which applies to CM prevention as well. With low-frequency phenomena like CM, the field cannot depend solely on prevention trials where randomization occurs at the level of the individual because selection and other factors severely limit generalizability. Consequently, prevalence measures such as CM rates, foster care placement rates, and CM injury rates, which are determined for whole catchment areas, take on much greater importance in the assessment of prevention outcomes.

State information systems pertinent to child maltreatment have undergone changes over several decades. The most prominent influence has been the Statewide Automated Child Welfare Information System (SACWIS), which is a federally funded, voluntary case management system. In 1993 the federal government authorized funding to assist states to develop SACWIS for the purpose of establishing an electronic case file for children and families served by the state child welfare agency. Over the years, states adopted SACWIS or developed their own variations. Each state has its own developmental history regarding changes in their respective data system.

3. Present case example

The context for the present case example is the state of South Carolina. The South Carolina child welfare system had a less integrated data system during the 2000–2006 time-period. In 2007 South Carolina initiated its own version of SACWIS called the Child Adult Protective Services System (CAPSS) in its Department of Social Services (SCDSS). The CAPSS later went through a second round of improvements three to four years after it was adopted. Presumably, there can be subtle changes of multiple origins over time which might alter data in the aggregate. As system changes are made, data sets are migrated to produce greater integration. There can be improvement of records retroactively as migration occurs and as refinement results from cross-dataset linkages. What is not known is the extent to which such system integration and data migration activities impact data sets in practice.

The execution thirteen years ago of an outcome study on CM prevention provided an excellent opportunity to examine data reproducibility for findings derived from child welfare system administrative records (Prinz, Sanders, Shapiro, Whitaker, & Lutzker, 2009; Prinz, Sanders, Shapiro, Whitaker, & Lutzker, 2016). This study used a place randomization design, where "place" (in this instance county) was the unit of random assignment to intervention or control conditions. The study found that community-wide implementation of a parenting and family support intervention (Triple P) produced positive impact on county-wide rates for substantiated CM cases, out-of-home placements, and hospital-treated CM injuries. The administrative data system, which was the source of the outcome measures at the time of the study, has since experienced routine changes that might have affected the data retroactively. The key reproducibility issue is whether and to what extent administrative record data re-examined a decade later for that same time period would yield comparable results to those based on data acquired at the time of the study. The re-examination focused on substantiated CM cases, which originated with the child protective services system, and on out-of-home placements, which originated with the foster care system, all of which were processed through the state's data repository agency.

4. Method

4.1. Data sources and information system

The original analyses are defined as the ones conducted in the published study (Prinz et al., 2009, 2016). Data for the original analyses were obtained through the South Carolina repository known at the time as the South Carolina Office of Research and Statistics (SCORS). Programming staff provided county-specific frequency counts per year for calendar years 1999–2005 for children birth to seven years (i.e., under age eight). The same types of data, compiled by current programming staff, were obtained in 2016 for the same time period from the repository currently known as the South Carolina Revenue and Fiscal Affairs (SCRFA) office (Health and Demographics Division), to conduct the reproducibility analyses.

4.2. Administrative records

4.2.1. Substantiated child maltreatment cases. County-level Child Protective Services caseworkers enter information on reported cases of substantiated CM into the CAPSS system maintained at the SCDSS state office. The SCDSS subsequently extracts this information to the data warehouse at SCORS. Substantiated CM cases involved any category of child abuse and/or neglect (all subsumed under the state system's term “child maltreatment”), including cases where more than one category was substantiated. For each county and year, SCORS staff programmers generated an aggregate frequency count for substantiated CM cases that involved all children birth to seven years (i.e., under age eight) but that was unduplicated so that no child was counted more than once in a calendar year. For the reproducibility study, SCRFA staff programmers generated new aggregate frequency counts meeting the same requirements.

4.2.2. Out-of-home placements. In a similar process to Child Protective Services, the foster care system caseworkers enter information on out-of-home placements into CAPSS. Extracts again are sent to SCORS. For each county and year, a SCORS staff programmer generated an aggregate frequency count for out-of-home placements for all children birth to seven years (i.e., under age eight) and that was unduplicated so that no child was counted more than once in a calendar year. For the reproducibility study, a SCRFA staff programmer generated new aggregate frequency counts for unduplicated out-of-home placements.

4.2.3. Population data for rate denominators. A prevalence rate for each of the two variables was calculated for each county per year by dividing the frequency count by an estimate of the number of children under eight years of age. In the original analyses, a static figure for each county was used across years based on census data. For the 2016 data release, the SCRFA was able to provide year-specific estimates for each county regarding the number of children who were under eight years of age based on a refined population demographics model not available at the time of the original data release.

4.3. Original study design and intervention

The original study tested the prevention of child-maltreatment-related outcomes from community-wide implementation of a multi-level system of parenting and family support (the Triple P—Positive Parenting Program). Triple P is a system of parenting interventions of varying intensities (from brief to more extended) and formats (individual, small group, large group, media communications), that was designed with population impact in mind (Sanders, 1999, 2008, 2012; Sanders, Kirby, Tellegen, & Day, 2014). Using a place-randomization design, eighteen mid-sized SC counties (each with a population between 50,000 and 175,000) were randomized to intervention or control conditions. In the intervention counties, several hundred practitioners in the existing workforce from many service sectors were trained to deliver Triple P to parents, while practitioners in the control counties engaged in services as usual without receiving Triple P professional training (and without even being contacted). The Triple P trained practitioners were located in a variety of sectors such as health care, preschool, childcare, elementary school, mental health, social services, and non-governmental community organizations. Calendar year 2005 was designated as the outcome year to gauge the impact from 24-month implementation of Triple P. The one-year baseline was calendar year 2003, and the five-year baseline was the average of calendar years 1999 through 2003. Additional details about the design, intervention, and study implementation are found in Prinz et al. (2009, 2016).

4.4. Analytic procedures

Because the original analyses were conducted with an averaged five-year baseline or a one-year baseline (Prinz et al., 2016), these two approaches were adopted for this reproducibility case study. An important initial step in the analysis is the establishment of baseline equivalence for the two outcome variables, which in this context refers to affirming that the intervention and control counties were comparable in the years prior to implementation of the intervention with respect to substantiated CM cases and out-of-home placements. Then, the main analyses involved analysis of covariance (ANCOVA) for each of the two outcome variables, controlling for either the five-year or the one-year baseline.

5. Results

5.1. Baseline equivalence

Baseline equivalence with the newly acquired data was examined again. At baseline, comparison of the intervention and control counties yielded no significant differences for substantiated CM cases at five-year baseline, $t\,(16) = 0.13$, $p\,(2T) = 0.90$, and one-year baseline, $t\,(16) = 0.40$, $p\,(2T) = 0.69$, and for out-of-home placements at five-year baseline, $t\,(16) = 0.26$, $p\,(2T) = 0.80$, and one-year baseline, $t\,(16) = 0.95$, $p\,(2T) = 0.35$. This baseline equivalence mirrored the baseline equivalence found in the original analyses.

Table 1
Comparison of Original and Updated Data and Analytic Results for Substantiated Child-Maltreatment Cases Per 1000 Children Under Age Eight.

Timeframe, Conditions, and Statistics		Original Data	Updated Data
Five-Year Baseline	Intervention Counties	10.82 (4.36)	11.93 (3.92)
	Control Counties	11.40 (6.75)	12.24 (6.17)
One-Year Baseline	Intervention Counties	10.62 (4.69)	10.39 (3.84)
	Control Counties	12.16 (7.38)	11.36 (6.09)
Outcome Year	Intervention Counties	10.86 (3.79)	9.13 (2.39)
	Control Counties	16.30 (9.13)	13.21 (6.64)
Outcome Analysis Using Five-Year Baseline			
	t-value for B_{cond}	$t(16) = 2.59$	$t(16) = 3.21$
	p-values (1T/2T)	$p = 0.02/.04$	$p = 0.003/.006$
	Effect size	Cohen's $d = 1.30$	Cohen's $d = 1.61$
Outcome Analysis Using One-Year Baseline			
	t-value for B_{cond}	$t(16) = 2.07$	$t(16) = 2.68$
	p-values (1T/2T)	$p = 0.028/.056$	$p = 0.009/.045$
	Effect size	Cohen's $d = 1.04$	Cohen's $d = 1.34$

Note: All data based on 9 intervention counties and 9 control counties.

Table 2
Comparison of Original and Updated Data and Analytic Results for Out-of-Home Placements Per 1000 Children Under Age Eight.

Timeframe, Conditions, and Statistics		Original Data	Updated Data
Five-Year Baseline	Intervention Counties	4.02 (1.59)	3.39 (1.56)
	Control Counties	3.76 (1.91)	3.19 (1.79)
One-Year Baseline	Intervention Counties	4.33 (1.51)	3.46 (1.42)
	Control Counties	3.25 (2.56)	2.77 (1.64)
Outcome Year	Intervention Counties	3.90 (2.11)	3.08 (2.10)
	Control Counties	4.64 (2.02)	3.91 (1.68)
Outcome Analysis Using Five-Year Baseline			
	t-value for B_{cond}	$t(16) = 1.74$	$t(16) = 2.11$
	p-values (1T/2T)	$p = 0.05/.10$	$p = 0.026/.052$
	Effect size	Cohen's $d = 0.87$	Cohen's $d = 1.06$
Outcome Analysis Using One-Year Baseline			
	t-value for B_{cond}	$t(16) = 2.20$	$t(16) = 2.19$
	p-values (1T/2T)	$p = 0.022/.044$	$p = 0.023/.045$
	Effect size	Cohen's $d = 1.10$	Cohen's $d = 1.10$

Note: All data based on 9 intervention counties and 9 control counties

5.2. Reproducibility analyses

In this case study, reproducibility analyses refer to the repeating of the original analytic methods using the revised data. For reproducibility analyses of substantiated CM cases, the ANCOVA result using the five-year baseline was: overall model $F(2,15) = 29.96$, $p < 0.001$, coefficient $B_{cond} = -3.810$ (*S.E.* 1.185), $t(16) = 3.21$, p (T/2T) = 0.003/.006. The ANCOVA result using the one-year baseline was: overall model $F(2,15) = 28.62$, $p < 0.001$, coefficient $B_{cond} = -3.252$ (*S.E.* 1.212), $t(16) = 2.68$, p (T/2T) = 0.009/.045. For substantiated CM cases, the results of the reproducibility and original analyses are summarized side by side in Table 1. The reproducibility analyses using both types of baselines yielded significant results in the preventive direction like those from the original analyses. However, effect sizes for the reproducibility analyses were substantially higher than those from the original analyses. With respect to the pattern of means, the original analyses reflected an increase in the mean rate for control counties compared with no increase for the intervention counties. In the reproducibility analyses, there was also an increase for the control counties but the intervention counties showed a reduction in the mean rate.

For reproducibility analyses of out-of-home placements, the ANCOVA result using the five-year baseline was: overall model $F(2,15) = 20.970$, $p < 0.001$, coefficient $B_{cond} = -1.030$ (*S.E.* 0.489), $t(16) = 2.11$, p (T/2T) = 0.026/.052. The ANCOVA result using the one-year baseline was: overall model $F(2,15) = 8.835$, $p = 0.003$, coefficient $B_{cond} = -1.447$ (*S.E.* 0.662), $t(16) = 2.19$, p (T/2T) = 0.023/.045. For out-of-home placements, the reproducibility and original analyses are summarized side by side in Table 2. The pattern of means for control versus intervention counties over time was comparable for the original and reproducibility analyses: mean rate went up for control counties and went down for intervention counties.

5.3. Post hoc exploration

Exploratory analyses were conducted to assess the extent to which changes in population estimates might be contributing to variance with the original data. The original frequencies for substantiated CM cases and out-of-home placements were used to compute rates based on the original versus revised population estimates. The difference for those rates regarding substantiated CM cases was an average of plus or minus 3.42% for 2003 and plus or minus 5.68% for 2005. The difference for those rates regarding out-of-home placements was an average of plus or minus 3.39% for 2003 and plus or minus 2.90% for 2005.

Using paired t-tests, rates based on original population estimates did not differ significantly from rates based on revised population estimates. For substantiated CM cases, t-tests for 2003 and 2005 respectively were: $t\ (17) = 0.28$, $p = 0.782$; $t\ (17) = 1.47$, $p = 0.16$. For out-of-home placements, t-tests for 2003 and 2005 respectively were: $t\ (17) = 0.33$, $p = 0.75$; $t\ (17) = 1.01$, $p - 0.33$.

Finally, correlations were computed to explore correspondence between the original and new data for raw frequencies without involving the population-denominator estimates. The database for these correlations was all 46 counties in South Carolina to produce more reliable estimates. Pearson correlation coefficients were computed for each of the years based on raw frequencies: (i.e., ignoring rates and population parameters): the mean correlation were 0.985 (range: 0.972–.995) for substantiated CM cases and 0.919 (range: 0.880–.942) for out-of-home placements.

6. Discussion

The field might assume that administrative data pertaining to the CM domain once recorded remain totally invariant over time. The results of this case study seem to suggest otherwise, namely that archival administrative data taken in the aggregate through a state's child welfare information system might undergo some subtle changes over time. However, data patterns and statistical inferences were essentially reproducible with only modest alteration after a lapse of ten years. For substantiated CM cases, the reproduced analyses produced higher effect sizes (Cohen's $d = 1.61$ and 1.34) than the original analyses (Cohen's $d = 1.30$ and 1.04), and reflected a reduction in mean rate for the intervention counties compared with an increase for the control counties. For out-of-home placements, the reproduced analyses produced similar effect sizes (Cohen's $d = 1.06$ and 1.10) as those in the original analyses (Cohen's $d = 0.87$ and 1.10).

Taken as a whole, this case study is a real-world example that provides support for the utility and viability of child-welfare system administrative records to evaluate preventive interventions. The CM prevention field is taking an increasing interest in testing and documenting impact of interventions on population prevalence rates but there are limited options. For example, reliance on Medicaid claims (i.e., government-based Medicaid health care coverage for economically disadvantaged individuals and families) leads to unacceptably low estimates of CM rates (Raghavan et al., 2015). Investigators need to count on child welfare system records of CM and out-of-home placements, as well as hospital and medical records associated with injuries and fatalities.

Place randomization outcome studies, as well as quasi-experimental outcome studies that match places without random assignment, are rare in the CM field and in all areas of prevention but are growing in number and importance. Most CM prevention research by necessity and practicality is focused at the level of the individual, which is essential for establishing and refining interventions. However, the ultimate goal is prevention, which necessitates demonstration of prevalence reduction. Methodologies have been developed to permit inferences at a population or aggregate level (Donner & Klar, 2000), which depends on reliable measures of the key constructs such as CM-related indicators.

The field is seeing an increasing use of archival administrative records to evaluate prevention. For example, a recent study documented the promising impact of a universal multi-session home visiting program in Durham County (NC) on CM cases and emergency room treatment (Dodge et al., 2014). A 15-year follow-up of universal implementation of Triple P when children were pre-schoolers demonstrated significantly lower rate of hospital emergency department visits in childhood and adolescence (Smith, 2015). County-wide implementation of multiple levels of Triple P over a five year period in Santa Cruz County (CA) produced meaningful reduction in substantiated CM (First 5 Santa Cruz County, 2016).

Prevention trials and program evaluation in the CM field can benefit from greater use of administrative records for documenting intervention impact over time. For community-wide interventions as well as policy-change initiatives, there are a number of advantages. Cost of using archival records is considerably lower than other forms of assessment and data collection. The nature of administrative record data is that it is a source that is independent from the intervention implementers for most studies or program evaluations. In many applications, administrative record data do not present the same kinds of missingness found with other types of data such as interviews, telephone surveys, or other individually administered measures. There are also disadvantages. For example, when a catchment area is too small or the time duration too brief, the rates of administrative CM-related data can be erratic and unreliable because of low base rates. Additionally, when prevalence rates derived from administrative records are the targets in a randomized prevention trial, investigators need to plan on having many geographic units to generate adequate statistical power.

This case study has some limitations. First, it was not possible to pinpoint precisely what affected the minor variations that were observed. For both the original and reproduced analyses, the programmers processed de-identified data in aggregate, which made it impossible to do a case by case comparison. Data differences between the two analyses were likely due to a combination of factors: (a) correction of missing and incorrect records in the year or two after recording; (b) data cleaning that occurred during migration and integration of datasets; and (c) more precise population estimates for the number of children under age eight in each county and year. However, the post hoc exploratory analyses were able to shed some light on the last source of variance. The more refined estimates for the number of children in the population of each county yielded rates that differed from the rates based on the original population estimates ranging from plus or minus 2.9% to plus or minus 5.7% depending on the variable and year. Nevertheless, comparison of the rates based on the original and revised population rates produced no significant within-county difference based on paired *t*-tests.

A second limitation is that the case example pertained to only one state information system and time frame. It is not clear how the observations apply to other states and time periods. SACWIS is operationalized throughout the U.S. but there are no doubt variations from state to state regarding each state's information system. And finally, this case study addressed the assessment of CM prevention, which might or might not have implications for other purposes of child welfare information systems such as documentation of needs, understanding interrelationships and trends among indicators and geographical variables, and gauging policy shifts. Related to this constraint, this case study also does not speak to the use of administrative child welfare data for research at the level of the individual (Green et al., 2015).

In conclusion, this case study demonstrated the reproducibility of archival administrative data in the context of a population outcome evaluation. Reliable population-prevalence measurement is an essential part of a comprehensive public health strategy aimed at the prevention of child maltreatment (Prinz, 2016).

Acknowledgments

The author would like to thank the staff at the Health and Demographics Division of the South Carolina Revenue and Fiscal Affairs Office for their excellent assistance; John Lutzker, Matthew Sanders, Cheri Shapiro, Diana Tester, and Daniel Whitaker for comments on an earlier draft; and David Finkelhor and Manuel Eisner for bringing data system considerations to the author's attention.

References

Baron, J. (2012). *Rigorous program evaluations on a budget: how low-cost randomized controlled trials are possible in many areas of social policy*. Coalition for Evidence-Based Policy.

Collins, F. S., & Tabak, L. A. (2014). NIH plans to enhance reproducibility. *Nature*, *505*(7485), 612–613.

Dodge, K. A., Goodman, W. B., Murphy, R. A., O'Donnell, K., Sato, J., & Guptill, S. (2014). Implementation and randomized controlled evaluation of universal postnatal nurse home visiting. *American Journal of Public Health*, *104*, s136–s143. http://dx.doi.org/10.2105/ajph.2013.301361

Donner, A., & Klar, N. (2000). *Design and analysis of cluster randomization trials in health research*. New York: Oxford University Press.

Feeney, L., Bauman, J., Chabrier, J., Mehra, G., & Woodford, M. (2015). *Using administrative data for randomized evaluations*. North America: J-PAL.

First 5 Santa Cruz County. (2016). *Strengthening families in Santa Cruz County: 5-year report 2010–2015. Technical report from First 5 Santa Cruz County*. the County of Santa Cruz Health Services Agency, and the County of Santa Cruz Human Services Department.

Gorski, D. (2016). Is there a reproducibility crisis in biomedical science? No, but there is a reproducibility problem. *Science-based Medicine*, https://www.sciencebasedmedicine.org/

Green, B. L., Ayoub, C., Bartlett, J. D., Furrer, C., Von Ende, A., Chazan-Cohen, R., et al. (2015). It's not as simple as it sounds: problems and solutions in accessing and using administrative child welfare data for evaluating the impact of early childhood interventions. *Children and Youth Services Review*, *57*, 40–49.

Haskins, R., & Margolis, G. (2015). *Show me the evidence: Obama's fight for rigor and results in social policy*. Washington, D.C: Brookings Institution.

Prinz, R. J., Sanders, M. R., Shapiro, C. J., Whitaker, D. J., & Lutzker, J. R. (2009). Population-based prevention of child maltreatment: The U.S. Triple P system population trial. *Prevention Science*, *10*, 1–12. http://dx.doi.org/10.1007/s11121-009-0123-3

Prinz, R. J., Sanders, M. R., Shapiro, C. J., Whitaker, D. J., & Lutzker, J. R. (2016). Addendum to: Population-based prevention of child maltreatment: The U.S. Triple P system population trial. *Prevention Science*, *17*, 410–416. http://dx.doi.org/10.1007/s11121-016-0631-x

Prinz, R. J. (2016). Parenting and family support within a broad child abuse prevention strategy: Child maltreatment prevention can benefit from public health strategies. *Child Abuse & Neglect*, *51*, 400–406.

Raghavan, R., Brown, D. S., Allaire, B. T., Garfield, L. D., Ross, R. E., & Hedeker, D. (2015). Challenges in using Medicaid claims to ascertain child maltreatment. *Child Maltreatment*, *20*, 83–91. http://dx.doi.org/10.1177/1077559514548316

Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The Triple P—Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review*, *34*, 337–357. http://dx.doi.org/10.1016/j.cpr.2014.04.003

Sanders, M. R. (1999). The Triple P—Positive Parenting Program: Towards an empirically validated multilevel parenting and family support strategy for the prevention of behavior and emotional problems in children. *Clinical Child and Family Psychology Review*, *2*, 71–90. http://dx.doi.org/10.1023/A:1021843613840

Sanders, M. R. (2008). The Triple P—Positive Parenting Program as a public health approach to strengthening parenting. *Journal of Family Psychology*, *22*, 506–517. http://dx.doi.org/10.1037/0893-3200.22.3.506

Sanders, M. R. (2012). Development, evaluation, and multinational dissemination of the Triple P–Positive Parenting Program. *Annual Review of Clinical Psychology*, *8*, 345–379. http://dx.doi.org/10.1146/annurev-clinpsy-032511-143104

Smith, G. (2015). *15-year follow-up of WA Triple P trial, collaboration for applied research and evaluation*. Western Australia: Telethon Institute for Child Health Research under contract with the Department of Health.

Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences*, *113*, 6454–6459.