Original Contribution

Organizational Designs for Achieving High Treatment Trial Enrollment: A Fuzzy-Set Analysis of the Community Clinical Oncology Program

By Bryan J. Weiner, PhD, Sara R. Jacobs, BS, Lori M. Minasian, MD, and Marjorie J. Good, RN, BSN, MPH, OCN

University of North Carolina at Chapel Hill, Chapel Hill, NC; and National Cancer Institute, Bethesda, MD

Abstract

Purpose: To examine the organizational design features that were consistently associated in 2010 with high levels of patient enrollment onto National Cancer Institute (NCI) cancer treatment trials among the oncology practices and hospitals participating in the NCI Community Clinical Oncology Program (CCOP).

Methods: Fuzzy-set qualitative comparative analysis was used to identify the recipes (ie, combinations of organizational design features) that CCOPs used to achieve high levels of patient enrollment onto NCI treatment trials in 2010. Four organizational design features were examined: number of open treatment trials with at least one patient enrolled, number of newly diagnosed patients with cancer, number of CCOP-affiliated physicians, and number of CCOP-affiliated hospitals or practices where patient enrollment could occur. Data were obtained from NCI data systems and CCOP grant progress reports.

Results: Two recipes were consistently associated with high levels of patient enrollment onto NCI treatment trials in 2010: having many open treatment trials and many new patients with cancer, and having many open treatment trials and many affiliated hospitals or practices. Together, these recipes accounted for nearly two thirds of CCOP membership in the high-performance set in 2010.

Conclusion: No single organizational design feature, by itself, was consistently associated with high levels of patient enrollment onto NCI treatment trials in 2010. Having a large menu of active treatment trials may be necessary to achieve high–patient enrollment performance, but this is not sufficient unless combined with either large patient volume or many participating sites.

Introduction

The local networks of oncology practices and community hospitals that participate in the National Cancer Institute (NCI) Community Clinical Oncology Program (CCOP) play an integral role in the cancer clinical research enterprise in the United States. Since 1983, these local networks (themselves called CCOPs) have enrolled approximately one third of all patients in NCI cancer treatment clinical trials and a large majority of participants in NCI cancer prevention and control clinical trials.[1] Not only have COPs helped the NCI advance the science of discovery by conducting research in clinical settings in which most people receive care, they have also accelerated the translation of research results into everyday clinical practice.[2-5]

For many CCOPs, 2010 proved to be a difficult year. Overall patient enrollment onto NCI treatment trials declined from the year before, and many CCOPs struggled to meet the expectations of the NCI for treatment trial enrollment. Several contributing factors seem to be at play. First, the NCI clinical trials menu has shrunk in the past 2 years. This decrease can be attributed in part to the consolidation taking place in the clinical cooperative group system in response to an influential yet critical report by the Institute of Medicine.[6] Second, NCI treatment trials are increasingly testing therapies applicable only to smaller subgroups of patients (eg, colon cancer with wild-type *KRAS*). For CCOPs, enrolling many patients is more challenging than it used to be, when trials were applicable to more patients. Finally, CCOPs are finding it increasingly difficult to enroll patients in the midst of a deep economic recession. Interviews with CCOP physicians and administrators suggest that patient volume has declined as patients put off medical care to manage other living expenses (Teal et al, manuscript submitted for publication). Patients are also hesitant to enroll, because they are worried about health insurance coverage and out-of-pocket costs.

Despite these challenges, some CCOPs exceeded NCI expectations for treatment trial enrollment. Because these challenging conditions will likely persist, CCOPs need to reconsider and possibly right size their organizational designs to adapt to what seems to be a new era in cancer clinical research. In this study, we employed a novel analytic method—fuzzy-set qualitative comparative analysis (fsQCA)—to identify the recipes (ie, combinations of organizational design features) that CCOPs used to achieve high levels of patient enrollment onto NCI treatment trials in 2010. To keep our results practical, we focused on design features that reflected key resource inputs or productive capabilities for CCOPs, did not depend directly on the level of funding received from the NCI, and could be modified by CCOP leadership.

Methods

Study Setting

This study focused on the NCI CCOP network, which has been described extensively elsewhere.[1,7-11] Briefly, the CCOP

network is a three-way partnership involving the NCI Division of Cancer Prevention (DCP), selected cancer centers and clinical cooperative groups (CCOP research bases), and community-based networks of hospitals and physician practices (CCOPs). The NCI DCP provides overall direction and funding for community hospitals and physicians to participate in clinical trials; CCOP research bases design clinical trials; and CCOPs assist with patient enrollment, data collection, and dissemination of study findings. As of December 2010, 47 CCOPs operated in 28 states and included 400 hospitals and more than 3,520 community physicians.[1] Minority-based CCOPs were excluded from analysis because they differ from CCOPs in their organizational structure and patient populations.

Study Design and Data Sources

This single-group, cross-sectional study included all 47 CCOPs in operation in 2010. We obtained data on CCOP 2010 patient enrollment onto NCI treatment trials and the 2010 treatment trial menu from the NCI CCOP, minority-based CCOP, and research base management system. We obtained data on CCOP volume of patients with cancer, affiliated physicians, and organizational structure from the progress reports that CCOPs submit to the NCI each February as a condition of CCOP program participation. The progress reports cover the 9-month period from June 2010 through February 2011. Although the time periods covered by the two data sources do not overlap perfectly, the organizational design features examined in this study exhibit only small fluctuations from year to year.

Measures

The study outcome was 12-month patient enrollment (ie, accrual) onto NCI treatment trials. Using substantive knowledge and prior research, we selected four organizational design features likely to drive CCOP treatment trial accrual,[7,8,11-18] Number of open treatment trials was measured as the number of NCI treatment trials for which the CCOP had at least one patient enrolled in 2010. Number of new patients with cancer was measured as the number of newly diagnosed patients with cancer seen in 2009 (the most recent year available for this study) at hospitals participating in a CCOP. Number of CCOP physicians was measured as the number of physicians registered as NCI investigators who could enroll patients onto NCI clinical trials in 2010. Number of CCOP components was measured as the number of hospitals, practices, or other sites in which patients could enroll onto NCI clinical trials in 2010.

Data Analysis

We employed fsQCA to identify the recipes that CCOPs used to achieve high levels of accrual to NCI treatment trials (ie, high performance). FsQCA uses a set-theoretic approach to identify the ways in which conditions (eg, CCOP organizational design features) combine in different ways to produce outcomes of interest.[19,20] We used fsQCA because we suspected that the recipes used by high-performing CCOPs were complex; their success resulted not from any single design feature but rather from the combination of several design features. In addition, we suspected that more than one recipe existed for achieving high performance. FsQCA is well suited for examining causal complexity (ie, multifactor recipes) and equifinality (ie, multiple recipes leading to the same outcome) in small- to medium-sized studies.[19,21-23] The Appendix (only only) provides a technical description of fsQCA.

Analysis proceeded in five steps. First, we transformed the study measures into set membership scores. To do so, we asked NCI DCP officials to use their expert knowledge to specify three values for each of the five study measures: full membership in the set of interest (definitely high-accrual performance), full nonmembership in the set of interest (definitely not high-accrual performance), and a cross-over point reflecting maximum ambiguity in membership in the set of interest (neither high-accrual performance nor not high-accrual performance). We used these values to transform study measures into scores ranging from 1.0, indicating full membership in the set, to 0.0, indicating full nonmembership in the set. For high-accrual performance, the full membership, cross-over, and full nonmembership values were 100, 70, and 50, respectively, because CCOPs must enroll a minimum of 50 patients onto NCI treatment trials to maintain CCOP funding. For many open treatment trials, the values were 50, 31, and 20, respectively. For many new patients with cancer, the values were 10,000, 5,000, and 3,000, respectively. For many CCOP physicians, the values were 100, 51, and 25, respectively. For many CCOP components, the values were 15, 12, and 4, respectively. Appendix Table A1 (online only) lists the fuzzy-set membership scores for each CCOP.

Second, we constructed a data matrix (known as a truth table) with 2^k rows, where k is the number of design features included in the analysis. Each row indicates a specific combination of organizational design features (ie, a recipe), with the full table listing all logically possible recipes ($N = 16$ in this study). We then sorted CCOPs into the rows of this truth table based on their fuzzy-set membership scores (Appendix, online only).

Third, we reduced the number of rows in the truth table based on two criteria. First, recipes had to exhibit at least one case.[19] Recipes with no cases were used in counterfactual analysis. Second, recipes had to exceed the minimum consistency threshold of 0.80 by an amount greater than could occur by chance.[19] Consistency refers to the degree to which cases displaying a given recipe also display the outcome of interest (Appendix, online only). Recipes had to meet both criteria to be considered in further analysis.

Fourth, we used Boolean algebra to eliminate logically redundant recipes (Appendix, online only). To illustrate, suppose some high-performing CCOPs have many open trials, see many patients with cancer, and have many components. Suppose other high-performing CCOPs have many open trials, see many patients with cancer, and do not have many components. These recipes can be logically reduced to CCOPs that have many open trials and see many patients with cancer, because the outcome is the same whether CCOPs have many components or not.

Finally, we performed a counterfactual analysis of recipes that lacked empirical cases. Counterfactual analysis asks, "If we had observed cases for recipes where we did not, how would the results change?" We limited our counterfactual analysis to easy counterfactuals, defined as situations in which a redundant condition (organizational design feature) is added to a set of conditions that already produces the outcome of interest.[19] Suppose high-performing CCOPs have many open trials, see many patients with cancer, and do not have many components. Suppose our study lacks empirical cases of CCOPs that have many open trials, see many patients with cancer, and do have many components. Substantive and theoretic knowledge suggest that having many components leads to high-accrual performance. An easy counterfactual analysis is that CCOPs that have many trials and see many patients are high performers whether or not they have many components. The recipe can be reduced to CCOPs that have many trials and see many patients, because adding another design feature (ie, many components) would not make a difference.

Results

In 2010, CCOPs accrued an average of 90 patients to NCI treatment trials (Table 1). CCOP enrollment ranged from a low of 17 patients to a high of 341 patients. CCOPs similarly varied widely in terms of the number of treatment trials to which they had accrued at least one patient, the number of newly diagnosed patients with cancer seen, the number of physicians affiliated with the CCOP, and the number of sites where patients could enroll onto NCI treatment trials.

Table 2 lists all 16 logically possible combinations of the organizational design features examined in this study and reports the number of CCOPs that exhibited each combination (ie, recipe). Five recipes met both the minimum frequency threshold of one case and the minimum consistency threshold of 0.80 by an amount greater than could occur by chance. These five recipes consistently led to high-enrollment performance in 2010. Not surprisingly, one winning recipe was to have many open treatment trials, see many patients with cancer, have many affiliated physicians, and have many sites where patients can enroll. However, this recipe was not the only recipe that consistently led to high performance.

These five recipes can be logically simplified into two recipes using Boolean algebra (Table 3). Having many open treatment trials and seeing many new patients with cancer is a recipe that consistently led to high-enrollment performance (recipe one). For CCOPs that did not have many affiliated physicians, having many open treatment trials and many component sites where patients can enroll onto clinical trials was also a recipe that consistently led to high performance (recipe two). The fact that having many open treatment trials is a condition shared by both recipes suggests this condition may be necessary for achieving high performance, whereas alone it is not sufficient.

Consistency scores are high for each recipe individually and for the two recipes together. High consistency scores mean that almost all of the CCOPs that followed these recipes exhibited high performance. Coverage scores indicate the percentage of cases that achieved high performance by using a given recipe, allowing one to evaluate the empirical relevance of a recipe. In terms of overall coverage, the set of recipes accounts for 63% of fuzzy membership in the outcome. The first recipe (ie, many trials, many patients) was the more empirically relevant recipe, both in absolute terms (raw coverage) and in relative terms (unique coverage). The smaller unique coverage values indicate that the two recipes overlap. Not only do the recipes share a condition (ie, many treatment trials), some CCOPs may be employing more than one recipe.

Table 1. CCOP Descriptive Statistics, 2010

Feature	No.	Mean	SD	Minimum	Maximum
Accrual to treatment trials*	47	90	67	17	341
No. of treatment trials	47	34	18	11	78
No. of patients with cancer	47	5,726	4,897	277	28,746
No. of CCOP physicians	47	48	40	11	209
No. of CCOP components	47	10	6	1	25

Abbreviations: CCOP, Community Clinical Oncology Program; SD, standard deviation.
* Accrual refers to the No. of patients enrolled onto National Cancer Institute–sponsored treatment trials.

Table 2. Truth Table Summarizing the Recipes for Achieving High-Accrual Performance

		Consistency			
Solution	No. of CCOPs	Value	Threshold	F-test	*P*
dtcm	15	0.26	0.80	45.50	.00
dtcM	0	0.56	0.80	5.34	.03
dtCm	1	0.75	0.80	0.42	.52
dtCM	0	0.75	0.80	0.30	.59
dTcm	7	0.74	0.80	0.49	.49
dTcM	1	0.88	0.80	3.16	.08
dTCm	**5**	**0.89**	**0.80**	**5.45**	**.02**
dTCM	0	0.84	0.80	0.11	.74
Dtcm	2	0.50	0.80	5.61	.02
DtcM	4	0.43	0.80	12.75	.00
DtCm	0	0.78	0.80	0.06	.80
DtCM	0	0.70	0.80	1.14	.29
DTcm	**2**	**0.93**	**0.80**	**16.25**	**.00**
DTcM	**1**	**0.93**	**0.80**	**12.04**	**.00**
DTCm	**2**	**0.94**	**0.80**	**18.72**	**.00**
DTCM	**7**	**0.93**	**0.80**	**5.77**	**.02**

NOTE. Bold font indicates recipes meeting both the minimum frequency threshold of one case and the minimum consistency threshold of 0.80 by an amount greater than could occur by chance. Upper-case letters indicate presence of organizational design feature (ie, many); lower-case letters indicate absence of organizational design feature (ie, not many).
Abbreviations: C, No. of component sites (eg, hospitals, practices) where patients can enroll onto NCI treatment trials; CCOP, Community Clinical Oncology Program; D, No. of newly diagnosed patients seen by CCOP; M, No. of physicians affiliated with the CCOP who can enroll patients onto NCI treatment trials; NCI, National Cancer Institute; T, No. of open NCI treatment trials with at least one patient enrolled.

Table 3. Simplified Recipes for Achieving High-Accrual Performance

Feature	Recipe One	Recipe Two
Many treatment trials	X	X
Many patients with cancer	X	
Many CCOP physicians		x
Many CCOP components		X
Raw coverage	0.50	0.33
Unique coverage	0.23	0.13
Consistency	0.94	0.90
Overall solution consistency	0.92	
Overall solution coverage	0.65	

NOTE. Upper-case X indicates causal condition present; lower-case x indicates causal condition absent.
Abbreviation: CCOP, Community Clinical Oncology Program.

Counterfactual Analysis

Counterfactual analysis of the five recipes that exhibited no empirical cases (Table 3) permits one to logically simplify the two recipes even further. Substantive and theoretic knowledge suggest that having many affiliated physicians should be associated with high enrollment performance. One could assume, therefore, that having many open treatment trials and many component sites should work just was well for CCOPs that have many affiliated physicians (not observed) as it does for CCOPs that do not have many affiliated physicians (observed). On the basis of this assumption, the second recipe could be simplified into many trials, many components. Counterfactual analysis produces no logically simpler recipe for the many trials, many patients recipe. Although the overall consistency and coverage scores for the two-recipe counterfactual analysis remain largely unchanged, the simplified recipes exhibit smaller unique coverage scores than they did before counterfactual analysis (data not shown). The simplified recipes exhibit greater overlap with each other because it is harder to identify CCOPs that are exclusively pursuing only one of the simplified recipes.

Sensitivity Tests

We reran the fsQCA model for high-enrollment performance twice to assess the sensitivity of our results. First, we increased the full membership, cross-over point, and full nonmembership values by 5% for all study measures. Then, we reset these values to 5% below the levels that NCI DCP officials specified. We observed the same recipes in both sensitivity tests that we observed in our original analysis. Our results seem robust to minor uncertainties in NCI DCP officials' substantive knowledge of what constitutes high-accrual performance, many open treatment trials, many CCOP-affiliated physicians, and many CCOP components.

Exploratory Analysis

Table 2 lists one recipe (ie, dTcM) that exceeded the consistency threshold of 0.80 but did not do so by an amount greater than could occur by chance. Had we relaxed the critical value for the probabilistic test for consistency from 0.05 to 0.08, we would have obtained a third simplified recipe for achieving high-enrollment performance. This simplified recipe included having many open treatment trials, many affiliated physicians, and not many component sites. Counterfactual analysis would have logically simplified this recipe further to many trials, many physicians.

Discussion

We sought to identify the organizational strategies (or recipes) that CCOPs used to achieve high levels of patient enrollment onto NCI treatment trials in 2010, a year that signaled major transitions in the NCI clinical trials research program. Two recipes were consistently associated with high levels of patient enrollment (> 100 enrolled patients). Both recipes were variations on the theme of size matters. One recipe involved having many open treatment trials (> 50) with at least one patient enrolled in 2010 and seeing many newly diagnosed patients with cancer (> 10,000). The other involved having many open treatment trials (> 50) with at least one patient enrolled in 2010 and having many hospitals or practices (> 15) where patients could enroll onto trials. The fact that having many open treatment trials with at least one patient enrolled in 2010 appeared in both recipes suggests that this organizational design feature was necessary, but not sufficient alone, to achieve high levels of patient enrollment. CCOPs needed to have more than just a large and active trial menu; they also needed to offer that large trial menu to as many patients as possible. Some high-performing CCOPs accomplished this by working with a few high-volume hospitals and practices, whereas others accomplished this by working with a large network of enrollment sites. Other recipes for achieving high performance might exist (eg, the third recipe revealed in exploratory analysis), but the two recipes that we identified in this study are the only ones among those that we studied that consistently led to high-enrollment performance in 2010 and accounted for nearly two thirds of CCOP membership in the high-enrollment group.

CCOPs receive NCI funding through peer-reviewed cooperative agreements that grant them flexibility to plan and allocate staff, office infrastructure, and other resources according to local needs. CCOP leaders could use our study findings to reconsider and possibly grow their current operations to make their clinical research programs maximally productive while ensuring that as many patients as possible have the opportunity to participate in clinical trials. Careful consideration should be given to both the costs and benefits of various expansion strategies, such as increasing the treatment trial menu, partnering with other high-volume institutions, and adding multiple enrollment sites. There is little value, for example, in opening trials onto which CCOP-affiliated physicians could not or would not enroll at least one patient in a year's time.

Several study limitations merit discussion. First, we employed data collected for program administration, not for research. Although we are confident about the accuracy of NCI treatment trial accrual data, we do not know how much noise exists in our organizational design features data. Data from CCOP grant progress reports have been used, however, in other

studies.[7,8,12-17,24] Second, we had imperfect measures of organizational design features. For example, we could only assess the number of treatment trials for which a CCOP had at least one patient enrolled. We do not know how many additional trials a CCOP had open that had no enrollment. Third, the recipes we identified might depend on or include unmeasured organizational design features. Finally, we limited our analysis to a single year (2010) that NCI DCP program officials deemed a watershed for the cancer clinical trials research program in the United States. Prospective replication with multiyear data would indicate the robustness of the recipes identified in our study.

In summary, no single organizational design feature, by itself, was consistently associated with high levels of patient enrollment onto NCI treatment trials in 2010. Having a large menu of treatment trials may be necessary to achieve high patient-enrollment performance, but it is not sufficient unless combined with either large cancer patient volume or many participating sites.

Accepted for publication on January 30, 2012.

Acknowledgment

Supported by Contract No. 050211-113011 from the National Cancer Institute.

Authors' Disclosures of Potential Conflicts of Interest

Although all authors completed the disclosure declaration, the following author(s) and/or an author's immediate family member(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: None **Consultant or Advisory Role:** None **Stock Ownership:** None **Honoraria:** None **Research Funding:** Bryan J. Weiner, National Cancer Institute; Sara R. Jacobs, National Cancer Institute **Expert Testimony:** None **Other Remuneration:** None

Author Contributions

Conception and design: Bryan J. Weiner, Sara R. Jacobs, Lori M. Minasian

Administrative support: Sara R. Jacobs

Collection and assembly of data: Sara R. Jacobs

Data analysis and interpretation: All authors

Manuscript writing: All authors

Final approval of manuscript: All authors

Corresponding author: Bryan J. Weiner, PhD, Cecil G. Sheps Center for Health Services Research, 725 Martin Luther King Jr Blvd, CB #7590, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-7590; e-mail: bryan_weiner@unc.edu.

DOI: 10.1200/JOP.2011.000507; published online ahead of print at jop.ascopubs.org on May 1, 2012.

References

1. Minasian LM, Carpenter WR, Weiner BJ, et al: Translating research into evidence-based practice: The National Cancer Institute Community Clinical Oncology Program. Cancer 116:4440-4449, 2010

2. Carpenter WR, Reeder-Hayes K, Bainbridge J, et al: The role of organizational affiliations and research networks in the diffusion of breast cancer treatment innovation. Med Care 49:172-179, 2011

3. Laliberte L, Fennell ML, Papandonatos G: The relationship of membership in research networks to compliance with treatment guidelines for early-stage breast cancer. Med Care 43:471-479, 2005

4. Warnecke RB, Johnson TP, Kaluzny AD, et al: The community clinical oncology program: Its effect on clinical practice. Jt Comm J Qual Improv 21:336-339, 1995

5. Johnson TP, Warnecke RB, Aitken MJ: Changing practice patterns, in Kaluzny AD, Warnecke RB (eds): Managing a Health Care Alliance: Improving Community Cancer Care. San Francisco, CA, Jossey-Bass Publishers, 1996, pp 105-128

6. Nass SJ, Moses HL, Mendelsohn J (eds): A National Cancer Clinical Trials System for the 21st Century: Reinvigorating the NCI Cooperative Group Program. Washington, DC, National Academies Press, 2010

7. Carpenter WR, Fortune-Greeley AK, Zullig LL, et al: Sustainability and performance of the National Cancer Institute's Community Clinical Oncology Program. Contemp Clin Trials 33:46-54, 2012

8. Carpenter WR, Weiner BJ, Kaluzny AD, et al: The effects of managed care and competition on community-based clinical research. Med Care 44:671-679, 2006

9. McKinney MM, Weiner BJ, Carpenter WR: Building community capacity to participate in cancer prevention research. Cancer Control 13:295-302, 2006

10. McKinney MM, Weiner BJ, Wang V: Recruiting participants to cancer prevention clinical trials: Lessons from successful community oncology networks. Oncol Nurs Forum 33:951-959, 2006

11. Kaluzny AD, Warnecke RB (eds): Managing a Health Care Alliance: Improving Community Cancer Care (ed 1). San Francisco, CA, Jossey-Bass Publishers, 1996

12. Kaluzny AD, Lacey LM, Warnecke R, et al: Predicting the performance of a strategic alliance: An analysis of the Community Clinical Oncology Program. Health Serv Res 28:159-182, 1993

13. Kaluzny AD, Lacey LM, Warnecke R, et al: Using a community cancer treatment trials network for cancer prevention and control research: Challenges and opportunities. Cancer Epidemiol Biomarkers Prev 3:261-269, 1994

14. Kaluzny AD, Lacey LM, Warnecke R, et al: Accrual of patients to randomized clinical trials: Factors affecting cancer prevention and control research. Int J Technol Assess Health Care 10:506-516, 1994

15. Kaluzny AD, Ricketts T 3rd, Warnecke R, et al: Evaluating organizational design to assure technology transfer: The case of the Community Clinical Oncology Program. J Natl Cancer Inst 81:1717-1725, 1989

16. Kaluzny AD, Warnecke RB, Lacey LM, et al: Cancer prevention and control within the National Cancer Institute's clinical trials network: Lessons from the Community Clinical Oncology Program. J Natl Cancer Inst 85:1807-1811, 1993

17. Klabunde C, Kaluzny A, Ford L: Community Clinical Oncology Program participation in the Breast Cancer Prevention Trial: Factors affecting accrual. Cancer Epidemiol Biomarkers Prev 4:783-789, 1995

18. Weiner BJ, McKinney MM, Carpenter WR: Adapting clinical trials networks to promote cancer prevention and control research. Cancer 106:180-187, 2006

19. Ragin CC: Redesigning Social Inquiry: Fuzzy Sets and Beyond. Chicago, IL, University of Chicago Press, 2008

20. Ragin CC: Fuzzy-Set Social Science. Chicago, IL, University of Chicago Press, 2000

21. Fiss PC: A set-theoretic approach to organizational configurations. Acad Manage Rev 32:1180-1198, 2007

22. Fiss PC: Building better causal theories: A fuzzy set approach to typologies in organization research. Acad Manage J 54:393-420, 2011

23. Vaisey S: Structure, culture, and community: The search for belonging in 50 urban communes. Am Sociol Rev 72:851-873, 2007

24. Lacey LM, Hynes DM, Kaluzny AD: Performance in quasi-firms: An example from the Community Clinical Oncology Program. J Health Hum Resour Adm 14:307-326, 1992

Appendix

Fuzzy-set qualitative comparative analysis (fsQCA) is a comparative analytic method grounded in set theory. Originally developed to support systematic comparisons across small to intermediate numbers of cases, fsQCA can also be applied in the analysis of large-N studies (Ragin CC: Chicago, IL, University of Chicago Press, 2008). FsQCA is a useful method for examining how causes and conditions combine in various ways to produce outcomes of interest. Although it is possible to examine such combinations through regression analysis, doing so requires the use of multiplicative interaction terms, which not only are cumbersome and difficult to interpret but also tend to be highly collinear with one another and with their component variables.

A set is simply a collection of objects, such as $X = \{a, b, c, d, e\}$. Membership in a set is determined by a rule or membership function: m_A. In classical set theory, any object belonging to a set takes on only two values, 0 or 1, and the mapping of the membership function over some space of objects Ξ is given as $m_A(x) : \Xi \rightarrow \{0, 1\}$. A fuzzy set is a set that is characterized by vagueness in the degree of set membership, wherein the numeric degree of membership can range from 0 to 1. Formally, the membership function over some space of objects Ξ mapping to the unit interval [0, 1] is given as $m_A(x) : \Xi \rightarrow [0, 1]$. Several methods exist for constructing the membership function for fuzzy sets (Ragin CC: Chicago, IL, University of Chicago Press, 2008; Fiss PC: Acad Manage J 54:393-420, 2011; Vaisey S: Am Sociol Rev 72:851-873, 2007; Verkuilen J: Sociol Methods Res 33:462-496, 2005). Following Ragin, we employed the direct method, in which experts use substantive and theoretic knowledge to map an underlying interval-scale variable into a membership scale by specifying values for three qualitative breakpoints: full membership, full nonmembership, and the cross-over point or point of maximum ambiguity about set membership. By specifying the direct method, FsQCA rescales the measures by using the cross-over point as an anchor from which deviation scores are calculated, taking the values of full membership and full nonmembership as the upper and lower bounds, respectively. Sensitivity testing with slightly different membership, nonmembership, and cross-over values is advisable to account for uncertainty in experts' substantive or theoretic knowledge (Skaaning S-E: Sociol Methods Res 40: 391-408, 2011).

Common operations with fuzzy sets include negation ("not"), union ("and/or"), intersection ("and"), and inclusion ("contained in"). Membership in fuzzy negation (symbol, ~) is defined as one minus the fuzzy membership in the set, that is, $m_X = 1 - X$. Membership in fuzzy union (symbol, ∪) is defined as $m_{X \cup Y} = \max(m_X, m_Y)$, that is, the maximum degree of membership in two sets (Ragin CC: Chicago, IL, University of Chicago Press, 2008; Smithson M, Verkuilen J: Thousand Oaks, CA, Sage Publications, 2006). Membership in fuzzy intersection (symbol, ∩) is defined as $m_{X \cap Y} = \min(m_X, m_Y)$, that is, the minimum degree of membership in two sets (Ragin CC: Chicago, IL, University of Chicago Press, 2008; Smithson M, Verkuilen J: Thousand Oaks, CA, Sage Publications, 2006). With fuzzy sets, objects have at least partial membership in all possible combinations (intersections) of sets. To assign cases to rows in the truth table, fsQCA assigns cases to the combination for which the case has a membership score greater than 0.5 (Ragin CC: Chicago, IL, University of Chicago Press, 2008). Such a score signals that the case is more in than out of the combination in question. Mathematically, a case can have only one membership score greater than 0.5 in all logically possible combinations of a given number of sets.

For set Y to include set X, the fuzzy membership scores for set X must consistently be lower than the fuzzy membership scores for set Y. A simple measure of inclusion is given by $\#(m_X \leq m_Y)/n$, or the proportion of cases that display this pattern of fuzzy membership relative to the total number of cases (Ragin CC: Chicago, IL, University of Chicago Press, 2008; Smithson M, Verkuilen J: Thousand Oaks, CA, Sage Publications, 2006). FsQCA uses a refined measure that assigns small penalties for minor inconsistencies and large penalties for major inconsistencies (Ragin CC: Chicago, IL, University of Chicago Press, 2008). Specifically, consistency $(X_i \leq Y_i) = \Sigma[\min(X_i, Y_i)] / \Sigma(X_i)$. Thus, partial credit is given for near misses, where membership scores in X are almost, but not quite, lower than membership scores in Y. Consistency indicates how closely a perfect subset relation is approximated. By establishing a benchmark for consistency scores, probabilistic tests can be employed to assess whether consistency (or the degree to which X is a subset of Y) is greater than could be expected by chance.

In fsQCA, the consistency of subset relations is used to assess the sufficiency of a causal condition or combination of causal conditions to produce an outcome. Sufficiency exists when all or nearly all cases that exhibit a causal condition (design features, in this study) or combination of causal conditions (recipes, in this study) exhibit the same outcome (high-accrual performance, in this study; Ragin CC: Chicago, IL, University of Chicago Press, 2008; Ragin CC: Chicago, IL, University of Chicago Press, 2000). Consistency highlights a key feature of subset relations: asymmetry. Y might fully include X, but X might not fully include Y. An alternative measure of consistency is used in fsQCA to assess the necessity of a causal condition or combination of casual conditions to produce an outcome. Necessity exists when all or nearly all instances of a specific outcome (high-accrual performance) exhibit the same causal condition (design feature) or combination of causal conditions (recipe; Ragin CC: Chicago, IL, University of Chicago Press, 2008; Ragin CC: Chicago, IL, University of Chicago Press, 2000). The consistency of a subset relation for a necessary condition is: consistency $(Y_i \leq X_i) = \Sigma[\min(X_i, Y_i] / \Sigma(Y_i)$. When fuzzy membership in set Y (the outcome) is consistently less than or equal to fuzzy membership in set X (a causal condition or combination of causal conditions), and consistency exceeds an established benchmark by an amount greater than could be expected by chance, then empirical evidence supports the proposition that X is a necessary condition for Y. The proposition must also be consistent with theoretic knowledge.

Coverage refers to the degree of overlap among two or more sets (Ragin CC: Chicago, IL, University of Chicago Press,

2008; Ragin CC: Chicago, IL, University of Chicago Press, 2000). In fsQCA, coverage is used to assess the degree to which a causal condition (design feature) or combination of causal conditions (recipe) covers or accounts for instances of an outcome (high-accrual performance). Thus, coverage gauges empirical relevance or importance. Coverage is measured as $(X_i \leq Y_i) = \Sigma[\min(X_i, Y_i] / \Sigma(Y_i)$, the same formula used to measure the consistency of a subset relation for a necessary condition. Ragin notes that this formula has a different meaning, and serves a different purpose, depending on the context in which it is used (Ragin CC: Chicago, IL, University of Chicago Press, 2008). In fsQCA, coverage is calculated only after establishing that a subset relation indicating sufficiency is consistent. There is little value in calculating the coverage of a recipe that is not consistently related to the outcome. When more than one recipe is consistently related to the outcome, the coverage of each recipe can overlap. As with the partitioning of explained variance in regression, the coverage of a recipe can be partitioned into shared coverage and unique coverage in fsQCA.

Boolean algebra is the algebra of sets (Ragin CC: Berkeley, CA, University of California Press, 1987). Boolean addition, for example, is equivalent to the set operation of union ("or"). Thus, in Boolean algebra, if $A + B = Z$ and $A = 1$ and $B = 1$, then $Z = 1$. Put another way, if $A = 1$ or $B = 1$ then $Z = 1$. Either result ($A = 1$ or $B = 1$) is sufficient for $Z = 1$. Boolean multiplication is equivalent to the set operation of interaction ("and"). Thus, in Boolean multiplication, the expression *Abc*, where upper-case letters indicate the presence of a condition, and lower-case letters indicate the absence of a condition, does not mean that the value of *A* (1) is multiplied by the value of *b* (0) and by the value of *c* (0), returning a result of 0. It simply means the presence of *A* is combined with the absence of *B* and the absence of *C*. It is a logical statement, not an arithmetic one. Thus the statement $Z = Abc + ABc + abC$ is a sum-of-products statement indicating that *Z* results when *Abc* occurs or *ABc* occurs or *abC* occurs. This statement can be simplified by the rule of minimization (Ragin CC: Berkeley, CA, University of California Press, 1987). Simply stated, "If two Boolean expressions differ in only one causal condition yet produce the same outcome, then the causal condition that distinguishes the two expressions can be considered irrelevant and can be removed to create a simpler, combined expression" (Ragin CC: Berkeley, CA, University of California Press, 1987, p 93). Because both *Abc* and *ABc* produce the same result *Z* yet differ only in the presence or absence of *B*, the two expressions can be replaced with a single, simpler expression *Ac*. In fsQCA, the combinations of causal conditions (recipes) that consistently lead to an outcome (high-accrual performance) are arrayed in a truth table and subjected to the minimization rule to logically reduce complex combinations of causal conditions (recipes) into simpler ones that still produce the same result (high-accrual performance).

Furthermore, in Boolean algebra, the expression *A* implies *Abc* because the set *A* embraces the subset *Abc* (Ragin CC: Berkeley, CA, University of California Press, 1987). Suppose *A* indicates the presence of open treatment trials, *B* indicates the presence of many CCOP-affiliated physicians, and *C* indicates the presence of many CCOP component sites. Set *A* embraces all CCOPs that have many open treatment trials. Set *Abc* embraces all CCOPs that have many open treatment trials and lack many CCOP-affiliated physicians and many CCOP components. Membership in *Abc* is included in set of *A*. Thus, *A* implies *Abc*.

In fsQCA, the notion of implication is used to reduce the complex expressions in the sum-of-products statements (called primitive expressions) found in the truth table into logically simpler expressions (called prime implicants; Ragin CC: Berkeley, CA, University of California Press, 1987). Often, this procedure produces more prime implicants than are needed to imply (ie, cover) all of the original primitive expressions in the sum-of-products statement. For example, the prime implicant *AB* implies the primitive terms *ABC* and *ABc*; however, these two primitive terms are also covered by *AC* and *Bc*, respectively. To determine whether *AB* is logically essential or logically redundant, a prime implicant chart can be created that maps the links between the prime implicants and the original primitive expressions. This chart can then be logically minimized to find the fewest prime implicants needed to cover all of the original primitive expressions. In fsQCA, the Quine-McCluskey algorithm is used to find the essential prime implicants needed to cover the primitive expressions in the sum-of-products statement (Ragin CC: Chicago, IL, University of Chicago Press, 2008). The suitability of the logically parsimonious solutions that these minimization procedures generate should be evaluated based on substantive and theoretic knowledge of the problem at hand.

Table A1. Fuzzy-Set Membership Scores for Treatment Trial Accrual and Four Conditions

CCOP Grant No.	Treatment Trials	Patients With Cancer	CCOP Physicians	CCOP Components	Treatment Accrual
1	0.69	0.86	0.92	0.88	0.94
2	0.98	0.49	0.65	0.32	1.00
3	0.65	0.42	0.04	1.00	0.60
4	0.78	0.36	0.05	1.00	1.00
5	0.62	0.05	0.17	0.05	0.11
6	0.00	0.39	0.08	0.10	0.00
7	0.13	0.00	0.07	0.13	0.79
8	0.54	0.88	0.31	0.25	0.67
9	0.89	0.88	0.63	0.13	0.95
10	0.69	0.03	0.01	0.03	0.52
11	0.03	0.55	0.22	0.13	0.04
12	0.03	0.01	0.02	0.07	0.00
13	0.00	0.07	0.02	0.02	0.00
14	0.87	0.80	0.73	0.95	0.99
15	1.00	1.00	1.00	1.00	1.00
16	0.01	0.01	0.02	0.03	0.00
17	0.31	0.01	0.07	0.05	0.03
18	0.06	0.97	0.80	0.32	0.46
19	0.16	0.98	0.75	0.41	0.11
20	0.85	0.02	0.06	0.99	0.62
21	0.69	0.63	0.41	1.00	0.79
22	0.75	0.50	0.79	1.00	0.07
23	1.00	0.50	0.07	1.00	1.00
24	0.78	0.07	0.09	0.03	0.18
25	0.13	0.01	0.02	0.07	0.01
26	0.04	0.59	0.22	0.25	0.00
27	0.01	0.06	0.02	0.05	0.00
28	1.00	0.62	0.47	0.07	1.00
29	0.03	0.99	0.99	0.32	0.05
30	0.01	0.15	0.07	0.07	0.46
31	0.31	0.09	0.03	0.32	0.16
32	0.94	0.45	0.20	0.13	0.99
33	0.10	0.28	0.20	0.05	0.02
34	0.94	0.07	0.17	0.10	1.00
35	0.99	1.00	1.00	1.00	1.00
36	0.02	0.02	0.01	0.05	0.00
37	0.02	0.04	0.07	0.05	0.94
38	0.99	1.00	0.95	0.73	1.00
39	0.43	0.17	0.17	0.88	0.79
40	0.83	0.35	0.33	1.00	0.67
41	0.97	0.03	0.06	0.05	1.00
42	0.83	0.99	0.99	0.99	0.94
43	0.01	0.04	0.02	0.02	0.00
44	0.65	0.18	0.03	0.88	0.75
45	0.10	0.66	0.93	0.03	0.05
46	0.98	0.00	0.04	0.25	0.86
47	0.13	0.01	0.01	0.18	0.04

NOTE. Calibration values are as follows: high accrual: fully in = 100, cross-over = 70, fully out = 50; many trials: fully in = 50; cross-over = 31; fully out = 20; many patients: fully in = 10,000; cross-over = 5,000; fully out = 3,000; many physicians: fully in = 100; cross-over = 51; fully out = 25; many components: fully in = 15; cross-over = 12; fully out = 4.
Abbreviation: CCOP, Community Clinical Oncology Program.