ORIGINAL ARTICLE

Do physicians within the same practice setting manage osteoporosis patients similarly? Implications for implementation research

J. R. Curtis · T. Arora · J. Xi · A. Silver · J. J. Allison ·
L. Chen · K. G. Saag · A. Schenck · A. O. Westfall ·
C. Colón-Emeric

Abstract

Summary Using data from long-term glucocorticoid users and long-term care residents, we evaluated osteoporosis prescribing patterns related to physician behavior and common practice settings. We found no significant clustering effect for common practice setting, suggesting that osteoporosis quality improvement (QI) efforts may be able to ignore this factor in designing QI interventions.

Introduction Patients' receipt of prescription therapies are significantly influenced by their physician's prescribing patterns. If physicians in the same practice setting influence one another's prescribing, evidence implementation interventions must consider targeting the practice as well as individual physicians to achieve maximal success.

Methods We examined receipt of osteoporosis treatment (OP Rx) from two prior evidence implementation studies: long-term glucocorticoid (GC) users and nursing home (NH) residents with prior fracture or osteoporosis. Common practice setting was defined as doctors practicing at the same address or in the same nursing home. Alternating logistic regression evaluated the relationship between OP Rx, common practice setting, and individual physician treatment patterns.

Results Among 6,281 GC users in 1,296 practices, the proportion receiving OP Rx in each practice was 6–100%. Among 779 NH residents in 66 nursing homes, the proportion in each NH receiving OP Rx was 0–100%. In both, there was no significant relationship between receipt of OP Rx and common practice setting after accounting for treatment pattern of individual physicians.

Conclusion Physicians practicing together were not more alike in prescribing osteoporosis medications than those in different practices. Osteoporosis quality improvement may be able to ignore common practice settings and maximize statistical power by targeting individual physicians.

Keywords Alternating logistic regression · Fracture · Glucocorticoid · Group randomized trial · Nursing home · Osteoporosis

J. R. Curtis (✉) · T. Arora · J. J. Allison · L. Chen · K. G. Saag
Center for Education and Research on Therapeutics,
University of Alabama at Birmingham,
Birmingham, AL 35294, USA
e-mail: jcurtis@uab.edu

J. Xi
Department of Sociology, University of Akron,
Akron, OH, USA

A. Silver · A. Schenck
The Carolinas Center for Medical Excellence,
Cary, NC, USA

K. G. Saag
Department of Epidemiology,
University of Alabama at Birmingham,
Birmingham, AL, USA

A. O. Westfall
Department of Biostatistics,
University of Alabama at Birmingham,
Birmingham, AL, USA

C. Colón-Emeric
Duke University and the Durham VA Medical Center GRECC,
Durham, NC, USA

Introduction

Osteoporosis is substantially under-recognized and under-treated. Even among the patients at highest osteoporosis risk,

such as long-term glucocorticoid users and persons with prior fracture, only approximately one-quarter to one-half of patients receive any form of prescription osteoporosis therapy [1–3]. One response to suboptimal management is to design quality improvement interventions to identify these types of patients and to improve their rate of receipt of medications that have been shown to attenuate the risk for fracture. In the USA, physicians may also be motivated to improve quality as a result of modest financial incentives to provide high quality of osteoporosis care for their patients with Medicare insurance [4]; it is anticipated that reimbursement will be even more tightly linked with quality in the future.

A goal of osteoporosis management is to assure that high-risk patients receive efficacious therapies to mitigate fracture risk, and the medications most strongly associated with risk reduction are available only by prescription. For this reason, the physician is a key component of quality improvement efforts. Moreover, physician prescribing behavior may be influenced by numerous factors including interactions with their peers, incentives, or restrictions on treatments that may vary by region, the local availability of certain resources (e.g., dual-energy X-ray absorptiometry testing), and other factors unique to the setting in which they practice. Thus, the influence of a common practice has traditionally been considered in designing implementation research interventions, and patients may be considered as being "nested" or "clustered" within their physician's practice. The effect of an individual physician's treatment proclivities, and similarities between patients treated in a common practice setting, has previously been demonstrated for receipt of osteoporosis therapies [5,6]. Statistical analyses are available to account for this type of clustering both at the physician level and at the physician group level.

Correspondingly, implementation research study designs also should consider the effect of common practice settings. When a significant clustering effect exists, statistical power for controlled trials of implementation research interventions is predominantly predicated on the number of units randomized rather than the number of patients analyzed [7]. For that reason, if physician groups or facilities are randomized rather than the individual physicians practicing within them, power will be lower than if individual physicians can be randomized. However, if there is substantial interaction between physicians within a common practice or group setting, randomization of individual physicians should be avoided as the intervention and control groups may "contaminate" one another. It is therefore vitally important to the planning of future evidence implementation studies to understand whether there is significant clustering at the level of the practice setting.

Using data from two previously published studies of osteoporosis management in long-term glucocorticoid users (GIOP dataset) and nursing home patients with known osteoporosis or prior fracture (SPOF dataset) [2,3,8,9], we evaluated whether there was significant similarity in osteoporosis management between physicians who practiced within the same outpatient group, or within the same nursing home. We tested the hypothesis that prior to any intervention and after controlling for patient factors (i.e., case mix) and individual physician clustering, there would be no effect of the physician group or nursing home setting on whether patients received screening or treatment for osteoporosis. The motivation for this hypothesis was to determine whether randomizing groups of physicians who practice together in a common setting, rather than randomizing individual physicians, was necessary. Randomizing groups of physicians would avoid contamination between intervention and control physicians but at the cost of sacrificing statistical power and raising other design issues.

Methods

Data sources

To evaluate the effect on osteoporosis management of group practice setting and an individual physician's proclivity to treat his or her patients similarly, we used datasets from two separate randomized, controlled studies of osteoporosis quality improvement interventions [2,3,8,9]. The GIOP dataset included information for 6,281 long-term glucocorticoid users enrolled in a large US healthcare organization. The information available in this data source consisted of administrative claims data and included complete information on demographics, comorbidities, prescription drug use, and health services utilization. The SPOF dataset included information for 779 ambulatory nursing home residents with a recent hip fracture or diagnosis of osteoporosis in Arizona and North Carolina [9]. These data were obtained through abstraction of nursing home medical records by trained nurses. Although both of these studies randomized physician practices or nursing homes to a quality improvement intervention, the data used for the current analysis were from the baseline data collection prior to any intervention to avoid any effect of the intervention itself.

Outcomes of interest

For patients represented in the GIOP dataset [2], the two outcomes of interest were receipt of bone mineral density testing and at least one filled prescription for an osteoporosis medication. For nursing home patients represented in

the SPOF dataset [3], we evaluated receipt of at least one prescription for an osteoporosis medication. In both datasets, the medications of interest included any oral bisphosphonate, teriparatide, calcitonin, and raloxifene. Prescriptions for hormone therapy were not evaluated in either dataset given diverse prescribing indications for this class of medications.

Definition of clustering

We evaluated the effect of physician groups within each of the two datasets. In the GIOP dataset, physician groups were defined as physicians who practiced at the same street address. In this dataset, we also evaluated the effect of physician specialty as recorded in the health plan's databases. For patients that were treated by more than one physician, the physician who prescribed the majority of the glucocorticoid prescriptions was identified as the physician responsible to assure appropriate osteoporosis care.

In the SPOF dataset, practice setting was defined as a nursing home; for the small number of physicians practicing at more than one nursing home within this dataset, they were represented at only the nursing home with the greatest number of their patients and their patients at other nursing homes were removed from the dataset. The physician designated as the nursing home attending on monthly order summaries was identified as the physician responsible to assure appropriate osteoporosis care.

Statistical analysis

Descriptive statistics were used to evaluate the mean proportion of patients receiving bone mineral density (BMD) testing or prescription osteoporosis therapies within each physician practice and nursing home; for this analysis, each physician group contributed one observation. These proportions were then evaluated at the level of the individual physicians, and the mean for each physician counted as one observation. To quantify the relative importance of the clustering effect at the physician group level and physician level, intraclass correlation coefficients (ICCs) were then calculated for each level. Assuming any two sub-units belonging to the same cluster are correlated equally, but the correlations between sub-units vary across clusters, generalized estimating equation method (GEE) quantifies the ICC by estimating the exchangeable working correlation. Extending the assumption to the three-level framework, we used the GEE approach to calculate ICCs by taking the residuals from the logistic regression models containing patient variables and estimating the ICC for physician groups as the correlation between any two residuals corresponding to patients from the same practice but not the same doctor. ICC for physicians were computed as the correlation between any two from the same doctor. Because

ICCs can also be explained as the proportion of the total outcome variation for corresponding cluster levels, after controlling for covariates that were identified in the original investigations [3,8], adjusted ICCs (i.e., residual ICCs) for the two cluster levels represent the fraction of total residual variation in the outcome attributable to the clustering.

Alternating logistic regression (ALR), a technique developed as an extension to generalized GEE was used to evaluate the association between OP Rx (receipt of prescription of non-hormone osteoporosis treatment) and physician group (within physician group cluster) by fitting a two-level-nested model where patients were clustered within the physician practice group. These results were re-evaluated after further accounting for the proclivity of individual physician to treat their patients similarly (within physician cluster) by fitting a three-level-nested model where patients were clustered within individual physicians, and physicians were clustered within physician groups. Both two-level-nested models and three-level-nested models were fitted with and without adjusting for the same covariates of interest. Age and sex were included as interaction terms in the fully adjusted models. The same analysis was then repeated for receipt of DXA as the outcome of interest. A parallel analysis was then conducted in the SPOF dataset, where the effects of clustering similarly was evaluated in two- and three-level models and accounted for the effects of the nursing homes, and then the simultaneous clustering effect of physicians practicing within those nursing homes. For this dataset, the outcome of interest was OP Rx. We also adjusted for a number of facility, physician group, and individual physician covariates.

Other statistical procedures besides ALR can account for clustered data. These other procedures include hierarchical linear models or random effects models implemented through procedures such as NLMIXED (SAS) or GLLAMM (STATA). These procedures require a different set of assumptions than ALR and are interpreted as the average change at the individual level and allow the effects of independent variables to be cluster-specific [10]. In contrast, ALR estimates a population-averaged result, which was felt to be more relevant for our analyses since we wished to predominantly focus on the magnitude of the clustering. Moreover, population-averaged results, such as those estimated by ALR, have been suggested to be more appropriate to account for correlated data in population-based studies [11]. However, in most studies, results computed with each of these methods are usually similar. Data analyses used SAS 9.1 (SAS Institute, Cary, NC, USA).

Power calculations were then performed to determine how many patients would be needed for a hypothetical randomized trial to detect a significant difference between an intervention vs. a control group on receipt of osteoporosis therapy [11]. A range of estimates was provided that

varied the treatment rate among the intervention group from 30% to 45% compared to a control group treatment rate of 25%. The number of persons that would be needed to detect these effect sizes was shown for an ICC of 0.00, representing simple randomization and no effect of clustering, up to an ICC of 0.09 which was at the higher end of the ICC range we observed.

We also plotted power as a function of the number of clusters randomized, holding the sample size constant [13]. The sample size was chosen to be approximately equal to the number of persons represented in the SPOF dataset. The purpose of this analysis was to evaluate the effect on statistical power of randomizing a smaller number of clusters (e.g., nursing home facilities, or physician groups), compared to a larger number of physicians. The former approach allows for the possibility of significant clustering within facilities and physician groups, whereas the latter approach yields better statistical power but requires the assumption that there is no meaningful clustering at the facility/physician group level.

Results

As shown in Table 1, in the GIOP dataset, there were 6,281 patients treated by 2,096 physicians practicing within 1,296 group practices. The mean proportion of GIOP patients receiving BMD testing and prescription osteoporosis medications across physician practices was 49% and 42%, respectively, but the range varied by more than tenfold, from less than 10% up to 100%. For the 779 nursing home residents treated by 246 physicians in 66 nursing homes, the mean receipt of prescription osteoporosis medications was 34% and varied between 0% and 100% across nursing homes.

Table 2 describes the effect of physician group and nursing home clustering effects on the receipt of BMD testing and prescription osteoporosis medications, both before and after adjusting for individual physician's proclivities regarding osteoporosis management. As shown in the two-level model, there was a strong and significant effect of clustering at the physician group and nursing home level. However, results from the three-level model demon-

Table 1 Proportion of patients screened or treated for osteoporosis, evaluating patients within physician groups and individual physicians

	GIOP dataset, n=6,281 patients	SPOF dataset, n=779 patients
Physician groups, n	1,296	66
Number of patients per group		
Median (mean)	4 (4.9)	11 (11.8)
Inter-quartile range	1–6	8–15
Range	1–50	2–26
Receipt of BMD testing, %		
Median (mean)	45 (49)	n/a
Inter-quartile range	25–64	
Range	8–100	
Receipt of OP Rx, %		
Median (mean)	33 (42)	30 (34)
Inter-quartile range	25–50	20–43
Range	6–100	0–100
Individual physicians, n	2,096	246
Number of patients per group		
Median (mean)	2 (3.0)	2 (3.2)
Inter-quartile range	1–4	1–4
Range	1–35	1–21
Receipt of BMD testing, %		
Median (mean)	50 (62)	n/a
Inter-quartile range	33–100	
Range	9–100	
Receipt of OP Rx, %		
Median (mean)	50 (55)	22 (34)
Inter-quartile range	25–50	0–60
Range	7–100	0–100

GIOP glucocorticoid induced osteoporosis, *SPOF* secondary prevention of fractures, *BMD* bone mineral density, *OP Rx* prescription osteoporosis medication, *n/a* not applicable

strated that this effect was significantly attenuated after accounting for clustering at the individual physician level. For receipt of osteoporosis medications and for receipt of BMD testing (in the GIOP dataset), the adjusted effect of physician group and nursing home was not significant.

Table 3 shows the effect of clustering on the number of persons that would be needed to recruit to a group-randomized trial after properly accounting for the similarity in osteoporosis management among patients treated by the same physician (or physician group). Compared to simple randomization, substantially greater numbers of patients would need to be recruited to such a trial throughout the range of ICCs that we observed.

The number of clusters randomized had a significant influence on study power, as shown in Fig. 1. With fewer than 200 clusters randomized, power was less than 80% and a strong inflection point was observed at approximately 100 clusters. With more than 400 clusters available to be randomized (as for the GIOP dataset), power appeared to

asymptomatically plateau. Using these results vis-a-vis the number of physician groups, nursing home facilities, and individual physicians from Table 1, there would have been a minimal gain in statistical power to have randomized physicians (rather than physician offices) in the GIOP study (1,296 physician groups, 2,096 individual physicians). In contrast, the gain in statistical power for the SPOF study (66 nursing home facilities, 246 individual physicians) would have been more substantial.

Discussion

In two separate datasets in which we evaluated long-term glucocorticoid users and nursing home patients with a prior fracture, we observed wide variability in osteoporosis management across outpatient physician practices and nursing homes, respectively. We observed that most of the clustering effects within these two types of practice settings

Table 2 Likelihood of patients treated within the same physician group or by the same physician to receive BMD testing or prescription osteoporosis medications

	Two-level model (physician group and patients)		Three-level model (physician group, physicians, and patients)	
	Crude OR (95% CI)	Adjusted[a] OR (95% CI)	Crude OR (95% CI)	Adjusted[a] OR (95% CI)
Glucocorticoid users, receipt of DXA				
Physician group	1.61 (1.30–1.98)	1.55 (1.16–2.10)	1.26 (1.09–1.44)	1.12 (0.98–1.28)
ICC		0.12		0.04
Individual physician	–	–	1.87 (1.39–2.52)	1.87 (1.24–2.84)
ICC				0.11
Glucocorticoid users, receipt of OP Rx				
Physician group	1.41 (1.20–1.67)	1.36 (1.18–1.57)	1.22 (0.88–1.69)	1.14 (0.87–1.49)
ICC		0.10		0.02
Individual physician	–	–	1.56 (1.34–1.82)	1.53 (1.33–1.77)
ICC				0.13
Nursing home residents, receipt of OP Rx				
Physician group	1.18 (1.02–1.36)	1.12 (1.00–1.25)	1.08 (0.91–1.30)	1.04 (0.95–1.14)
ICC		0.04		0.01
Individual physician	–	–	1.32 (1.04–1.67)	1.22 (0.98–1.51)
ICC				0.05

These pairwise odds ratios (*ORs*) are referent to patients treated by different physicians and were computed using alternating logistic regression (*ALR*)

ICC intraclass correlation coefficient

[a] Results for long-term glucocorticoid users were adjusted for the number of physicians in the group; physician covariates including specialty, age, gender, and years since obtaining medical degree; and patient covariates including sex, age, prior fracture, number of provider visits, number of comorbid conditions, and prednisone use (cumulative dose and new vs. prevalent use). Results for the nursing home residents were adjusted for facility level covariates including the number of beds in the facility (i.e., size), for-profit status, rural (vs. urban), and proportion of Medicare patients; and patient covariates including gender, age, race, whether ambulatory, history of falling, cognitive impairment, GERD, dysphagia, esophagitis or peptic ulcer, breast cancer, thromboembolic disease, tobacco use, and alcohol abuse

Table 3 The number of persons needed to adequately power a group randomized trial of osteoporosis treatment depends on the magnitude of clustering

Hypothesized proportion of patients treated in an intervention group (%)	Intraclass correlation coefficient				
	0.00 (i.e., no clustering)	0.01	0.03	0.06	0.09
30	2,504	2,610	2,810	3,110	3,410
35	656	690	740	820	900
45	176	190	200	220	240

Data shown are the number of persons that need to be recruited in order to detect a significant difference with 80% power between the hypothesized intervention group treatment rate and a 25% control group treatment rate, mean cluster size=5. For simplification, one level of clustering is assumed. Calculations are from [12]

were due to the treatment proclivities of individual physicians rather than shared practice patterns related to the practice setting that might influence the physicians.

Our findings have implications for osteoporosis implementation research efforts in which interventions need to consider whether to target physicians, the healthcare environment including group practice setting, and/or patients. Although ideally an intervention would target all of these, we showed that the effect of osteoporosis management at the individual physician level outweighs the effect of the setting in which physicians practice together. Largely because of fear for contamination between intervention and control group physicians, the two trials

represented in this analysis randomized at the practice setting rather than individual physicians [8,9]. This had the effect of reducing power because fewer units were randomized. The effects on study power as a function of the number of units randomized shown in Fig. 1 demonstrate that in the range of cluster sizes relevant for the SPOF study, the decision regarding which level to randomize was of high importance. Moreover, at least prior to intervention, our results indicate that the groups in which physicians practice appear to have only a small effect on other physicians' osteoporosis treatment patterns. Indeed, the adjusted odds ratios of the effect of the practice setting ranged from 1.04 to 1.14, which might be argued to be clinically irrelevant. However, although these data are suggestive that the effect of clustering at the practice setting level is small, contamination still might occur in the context of a potent evidence implementation intervention. This observational study cannot fully discount this possibility, and cluster randomization by facility should still be conducted until the potential for contamination has been further evaluated. Ideally, this could take place in the context of a study that randomized in one arm at the facility level, and in the other arm, at the provider level. However, this approach may or not be feasible.

Our results are consistent with a prior report that evaluated treatment patterns among 1,973 predominantly postmenopausal women patients treated by 435 primary care physicians practicing in the northeastern USA [14]. In that study, the magnitude of the clustering observed among physicians (ICC=0.03–0.04) was similar to our results where we observed ICCs ranging from 0.04 to 0.12

Fig. 1 Power based upon randomizing varying numbers of clusters, holding sample size constant (n=792 patients)
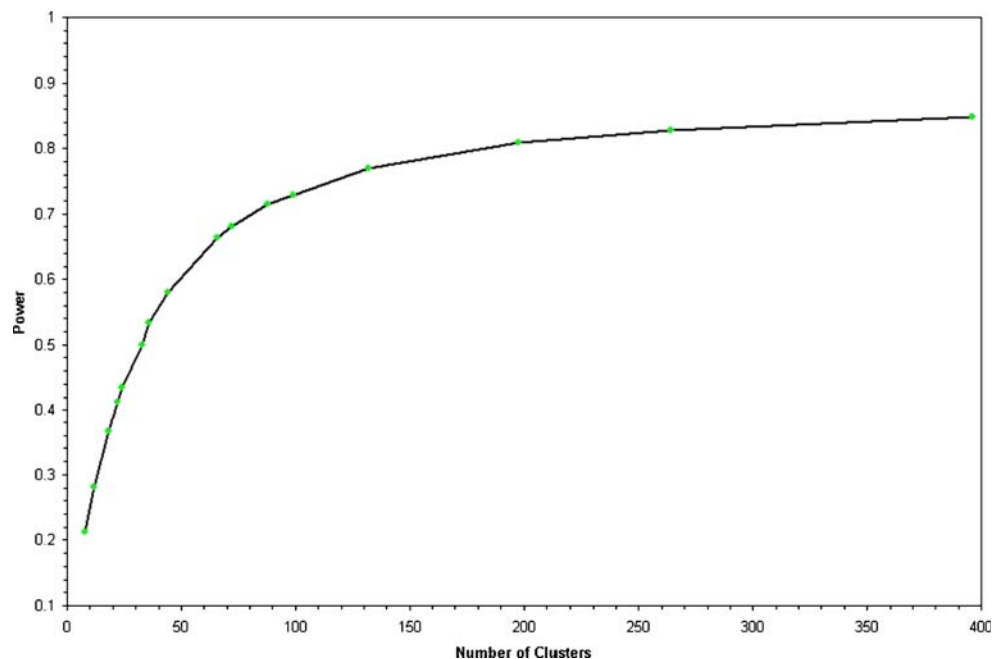
(depending on the practice setting and the outcome of BMD testing or receipt of prescription osteoporosis therapy). Our work extends those observations by focusing on two populations at high risk for fractures, long-term glucocorticoid users and nursing home residents with prior fracture. Additionally, we considered the effect of both clustering at the physician level and also the common practice setting (i.e., the outpatient clinic and the nursing home). We also were able to evaluate and control for the specialty of the physicians and a number of other physician, physician group, and facility covariates.

The strengths of our work include demonstrating consistent results in two separate datasets with unique patient populations that both considered the same osteoporosis endpoint. These two populations, long-term glucocorticoid users and older patients with prior fracture, represent individuals for which the strongest evidence and most robust osteoporosis management guidance exist. Additionally, this work should help guide future osteoporosis and other chronic disease quality improvement that use a group randomized trial design with multiple levels of clustering. As a potential limitation, we did not have information about whether the outpatient physician practices had DXA scanners in their office, which might account for some physician group clustering for the BMD testing outcome. We also recognize that these long-term glucocorticoid users were enrolled in a large commercial US healthcare organization and the nursing homes studied were from only two US states, and the generalizability of our findings may not extend to other populations.

In conclusion, we observed that patients receiving care in the same outpatient physician practices and nursing homes were significantly more likely to receive similar care than patients in different physician practices and nursing homes. Most of this effect was a result of individual physicians' treatment patterns rather than the shared practice setting. Although osteoporosis implementation research interventions are most likely to be successful if they can target all facets of the healthcare environment, the treatment patterns of individual physicians appear to outweigh the effects of the common settings in which they practice. The number of units randomized also was important in determining the study power, suggesting that if there is no effect of clustering at the physician group level, randomizing at the physician level would be preferable to maximize study power. This decision would require the assumption that these observational results apply in the context of a randomized evidence implementation intervention. In future studies, and depending on the number of physicians and physician groups available to randomize, baseline data on these hierarchical relationships are likely to be useful in the design phase to maximize study power.

Conflicts of interest None.

Disclosures for unrelated projects JRC: Consulting: Roche, UCB, Procter & Gamble; speakers bureau: Merck, Procter & Gamble, Eli Lilly, Roche, Novartis; research grants: Merck, Procter & Gamble, Eli Lilly, Amgen, and Novartis.

References

1. National Committee for Quality Assurance: State of Health Care Quality. 2006
2. Curtis JR, Westfall AO, Allison JJ et al (2005) Longitudinal patterns in the prevention of osteoporosis in glucocorticoid-treated patients. Arthritis Rheum 52(8):2485–2494
3. Colon-Emeric C, Lyles KW, Levine DA et al (2007) Prevalence and predictors of osteoporosis treatment in nursing home residents with known osteoporosis or recent fracture. Osteoporos Int 18(4):553–559
4. Physician Voluntary Reporting Program. Centers for Medicare & Medicaid Services; 2007
5. Colon-Emeric CS, Casebeer L, Saag K et al (2005) Barriers to providing osteoporosis care in skilled nursing facilities: perceptions of medical directors and directors of nursing. J Am Med Dir Assoc 6(3 Suppl):S61–S66
6. Brookhart MA, Solomon DH, Wang P, Glynn RJ, Avorn J, Schneeweiss S (2006) Explained variation in a model of therapeutic decision making is partitioned across patient, physician, and clinic factors. J Clin Epidemiol 59(1):18–25
7. Murray DM (1998) Design and analysis of group randomized trials . Oxford University Press, Oxford
8. Curtis JR, Westfall AO, Allison J et al (2007) Challenges in improving the quality of osteoporosis care for long-term glucocorticoid users: a prospective randomized trial. Arch Intern Med 167(6):591–596
9. Colon-Emeric CS, Lyles KW, House P et al (2007) Randomized trial to improve fracture prevention in nursing home residents. Am J Med 120(10):886–892
10. Preisser JS, Arcury TA, Quandt SA (2003) Detecting patterns of occupational illness clustering with alternating logistic regressions applied to longitudinal data. Am J Epidemiol 158(5):495–501
11. Feng Z, Diehr P, Peterson A, McLerran D (2001) Selected statistical issues in group randomized trials. Annu Rev Public Health 22:167–187
12. Snijders TAB (2005) Power and sample size in multilevel linear models. In: Everitt BS, Howell DC (eds) Encyclopedia of statistics in behavioral science, vol 3. Chicester, Wiley, pp 1570–1573
13. Campbell MK, Thomson S, Ramsay CR, MacLennan GS, Grimshaw JM (2004) Sample size calculator for cluster randomized trials. Comput Biol Med 34(2):113–125
14. Glynn RJ, Brookhart MA, Stedman M, Avorn J, Solomon DH (2007) Design of cluster-randomized trials of quality improvement interventions aimed at medical care providers. Med Care 45(10 Supl 2):S38–S43