Data-Driven Clinical and Cost Pathways for Chronic Care Delivery

Yiye Zhang, PhD, and Rema Padman, PhD

ABSTRACT

OBJECTIVES: This study illustrates a systematic methodology to embed medical costs into the exact flow of clinical events associated with chronic care delivery. We summarized and visualized the results using clinical and cost data, with the goal of empowering patients and care providers with actionable information as they navigate through a multitude of clinical events and medical expenses.

STUDY DESIGN: We analyzed the electronic health records (EHRs) and medication cost data of 288 patients from 2009 to 2011, whose initial diagnoses included chronic kidney disease stage 3, hypertension, and diabetes.

METHODS: We developed chronological pathways of care and costs for each patient from EHR and medication cost data. Using a data-driven method called clinical pathway (CP) learning, which leverages statistical machine-learning algorithms, we categorized patients into clinically similar subgroups based on progressing clinical complexity and associated care needs. The CP-based subgroups were compared against cost-based subgroups stratified by quartiles of total medication costs, and visualized via pathways that are color-coded by costs.

RESULTS: Our methods identified 3 CP-based, and 4 cost-based, patient subgroups. Two sets of subgroups from each approach indicated some clinical similarity in terms of average statistics, such as number of diagnoses and medication needs. However, the CP-based subgroups displayed significant variation in costs; conversely, large differences in clinical needs were observed among cost-based subgroups.

CONCLUSIONS: This study demonstrates that CPs extracted from EHRs can be enhanced with appropriate cost information to potentially provide detailed visibility into the variability and inconsistencies in current best practices for chronic care delivery.

Am J Manag Care. 2016;22(12):816-820

Patient care journeys span the entire continuum of clinical and nonclinical experiences. Navigating through a multitude of these events is a mystifying and daunting experience for most patients, and particularly so for patients with multiple chronic conditions (MCCs).[1] Despite being the central figure in their scenarios, research indicates that patients hold little power for steering in an appropriate direction through these passages.[2] Although hospitals and individual providers recognize the need to guide patients through their treatments, often such guidance follows a top-down approach and does not actively involve patients in the decision-making process.[3] However, with personalized and precision medicine gaining increasing attention, and with more patients and providers participating in the value-based payment models that emphasize patient satisfaction,[4] healthcare stakeholders are seeking to facilitate shared decision making by incorporating not only treatment efficacy and clinical factors, but also—equally important—patients' personal preferences and financial limitations. In fact, multiple research studies have found that involving patients in the decision-making process can lead to higher satisfaction, lower medical costs, and better health outcomes, among other benefits.[5,6] Prior research also shows that patients' knowledge of pricing information is associated with lower medical costs for common medical services.[7]

With medical cost being such an opaque subject, providers also may not have the best guidance strategy for the treatments that they offer to their patients.[8,9] Given these challenges, this study proposes a new approach to incorporate medical costs explicitly in the chronologically ordered, clinical pathways (CPs) of patient experiences. We especially focused on MCC patients, who are high utilizers of medical services and often participate in innovative payment programs.[4] We illustrated this approach using a cost-centered perspective as well as a clinically focused perspective to show alignment in some subgroups and significant variations in others, in the categorization of pathways and patient subgroups under these 2 differing views. The long-term goal of this research is to eventually achieve accurate predictions of anticipated future events and costs following different clinical and cost pathways for improved shared decision making, and, subsequently, identify appropriate ranges of cost for targeted clinical pathways within a patient population.

TAKE-AWAY POINTS

This preliminary study extracts patterns and variations in medical spending associated with the flow of clinical events in the management of patients with multiple chronic conditions.

- It illustrates a systematic, generalizable methodology for embedding cost information into the clinical pathways of patient cohorts using electronic health record and cost data.
- It exposes variations in costs among clinically similar patients, and variations in clinical complexity among patients with similar costs, by comparing patient cohorts categorized by their clinical pathways and total medical spending, respectively.
- Future extensions will examine appropriate incorporation of data-driven evidence into shared decision making and innovative cost analyses.

Background

CPs, commonly referred to as treatment plans that are time- and task-oriented, describe the essential steps associated with the expected clinical course in diagnosing and treating patients.[10] The goals in using CPs are to optimize treatment efficiency, patient outcomes, and medical spending.[10] CPs are used widely in the United States and are expected to have an even larger influence on healthcare delivery, quality, and patient outcomes, as bundled or episode-based payments gain momentum. Furthermore, using information from CPs in the operational setting has the potential to optimize not only the efficacy of treatments, but also their costs to patients, providers, and society.

Despite these anticipated benefits, to the best of our knowledge, there is a tremendous gap in the resources that will allow clinicians and patients to incorporate best practices for healthcare delivery in the form of CPs. Therefore, we aimed to begin bridging this gap by developing a data-driven approach that would take advantage of the large amount of detailed, patient-level, clinical information captured in electronic health records (EHRs), coupled with cost data that document the medical costs associated with patients' clinical activities. This preliminary study applies an advanced, systematic, and generalizable methodology to embed medical costs, specifically, costs of medications, into the exact flow of clinical events associated with the complexities of chronic care delivery. Combined pathways of clinical and cost information are summarized and visualized for patient subgroups following common CPs, and compared against simple cost-based segmentation of the patient population.

METHODS

CP learning refers to an emerging area of data-driven research to mine the common and rare patterns in the co-progression of 1 or more clinical events, depending on the availability of granular data and types of research questions, using statistical machine-learning methods.[11] Our previous studies developed and applied these methods to chronic kidney disease (CKD) to identify CPs containing: a) encounters; b) diagnoses as *International Classification of Diseases, Ninth Revision*, *Clinical Modification* (*ICD-9-CM*) codes; c) procedures as Current Procedural Terminology (CPT) codes; and d) prescriptions of medications summarized by drug classes.[12,13] This study extends these methods that combine statistical machine learning and data visualization by specifically incorporating costs associated with medications. Developing such detailed visibility into the co-progression of treatments and diseases from actual practice data, rather than abstract clinical guidelines, along with the costs associated with the events—medications, in this instance—is potentially a useful and novel approach to investigate.

We applied 2 approaches in identifying CPs with cost data. The clinically focused approach used our CP-learning algorithm to detect the differences in progression of clinical factors, but not costs, in order to divide patients into clinically similar subgroups. The costs of medications were subsequently computed for these subgroups. The cost-centered approach, which is analytically simpler, divided patients into similar subgroups based solely on the total amount of estimated co-pays on medications. Generally, the clinically focused approach categorizes patients by pathways of care, whereas the cost-centered approach does so based on the total cost of care.

Sample Patients

We collaborated with a community nephrology practice in western Pennsylvania that specializes in the management of CKD. CKD patients—mostly an elderly and vulnerable population—commonly suffer from a number of serious comorbidities and complications as the disease progresses to end-stage renal disease (ESRD).[14] Many CKD patients are Medicare beneficiaries, and the United States Renal Data System reports that CKD patients incur nearly twice the cost for their care, on average, compared with the non-CKD Medicare patient.[14] In addition, characteristics of the disease have resulted in limited high-quality randomized clinical trials, further exacerbating the challenges in care delivery.[15] Hence, improved management of this population using the approach illustrated in this study may generate multiple benefits, such as reduced costs and increased patient satisfaction.

Data

We obtained detailed clinical data, over the time period from 2009 to 2011, from the EHR extracts of the nephrology practice. We identified 288 patients diagnosed with only CKD stage 3, diabetes, and hypertension, and with no other complications, at the beginning of the 2-year period. The gender ratio of female to male was 0.41 to 0.59, and the average age of patients was 73.4 years (standard deviation [SD] = 10). Race represented by the study patients included Caucasian (94.4%), African American (4.8%), and other (0.8%). Over the 2-year

TABLE. Descriptive Statistics of Subgroups by Spending and Clinical Complexity, Over a 2-Year Period

Subgroup	Category	Categorize by Cost				Categorize by CP		
		High Spending	Medium Spending	Low Spending	Zero Spending	High Complexity	Medium Complexity	Low Complexity
	Patients, n	72	72	85	59	106	92	90
Demographics	Avg age, years (SD)	72.2 (9.0)	71.4 (11.3)	75.0 (9.3)	74.9 (9.8)	73.2 (9.9)	74.7 (9.3)	72.2 (10.6)
	African American, %	5.6%	1.4%	5.9%	6.8%	2.8%	7.6%	4.4%
	Female, %	37.5%	41.7%	40.0%	45.8%	39.6%	39.1%	44.4%
Cost, $	Avg	3826.4	274.6	25.2	0.1	2512.7	272.5	66.8
	SD	3215.0	199.6	21.9	0.3	2705.6	1960.3	135.0
	CV	0.84	0.73	0.87	3.00	1.08	7.19	2.02
	Max	18,801.4	699.8	78.6	1.8	10,993.5	18,801.4	590.3
	Min	722.6	78.9	2	0	12	0	0
Clinical	Avg number of unique diagnoses (SD)	6.7 (2.04)	5.6 (1.65)	5 (1.55)	4.8 (1.49)	6.2 (1.92)	5.3 (1.64)	5 (1.67)
	Patients with CKD progression, %	29.20%	16.70%	7.10%	8.50%	23.60%	14.10%	6.70%
Service utilization	Avg number of visits (SD)	7.0 (2.64)	6.2 (2.24)	5.4 (2.04)	4.9 (2.18)	6.4 (2.38)	5.4 (2.65)	5.9 (2.06)
	Avg number of unique drug class prescriptions (SD)	16.7 (5.97)	9.7 (5.16)	3.6 (2.66)	0.4 (0.80)	15.2 (6.22)	3.6 (3.66)	3.2 (3.63)

Avg indicates average; CKD, chronic kidney disease; CP, clinical pathway; CV, coefficient of variation; max, maximum; min, minimum; SD, standard deviation.

period, patients had, on average, 5.5 (SD = 2.1) office visits, 0.4 (SD = 1.0) hospitalizations, and 0.4 (SD = 0.2) education sessions, respectively.

CPs were created based on all records within 2 years since the initial visit, such that we could track the co-progression of clinical events and associated costs within a fixed time interval. Analysis includes the following patient care information: a) encounter types: office visit, hospitalization, education sessions; b) diagnoses as *ICD-9-CM* codes: CKD stages 1 to 5, ESRD, hypertension, diabetes, acute kidney injury, hyperparathyroidism, anemia, proteinuria, hyperkalemia, acidosis, hyperphosphatemia, glomerulonephritis, urinary obstruction, volume depletion, rhabdomyolysis; c) procedures as CPT codes: renal Doppler and ultrasound; and d) all medications taken by patients, excluding over-the-counter medications.

In the clinically focused approach, medications are summarized by drug classes for the purpose of patient subgrouping. In the cost-centered approach, total costs are calculated for each medication as is. Therefore, it is possible for 2 patients with similar clinical conditions and needs to be categorized into different levels, such as medium for complexity and high for spending, because CP-based subgroups are based on diagnoses and drug classes, whereas cost-based subgroups are based on costs from co-pays of each medication brand. For the sake of simplicity in illustrating our approach, the analysis in this paper examined only the cost of medications in the form of patient co-pays. We manually obtained the estimated co-pay of each medication from GoodRx that provides information on prescription prices under different insurance plans. To do so, we assumed that all patients were: Medicare beneficiaries, under AARP Medicare Complete Plan 1; residing in zip code 15201 of Pittsburgh, Pennsylvania; being prescribed medication dosages at the lowest appropriate level; and purchasing 30-day medication supplies.

RESULTS

The total estimated medication co-pays over 2 years had a wide range, from \$0 to over \$18,800, with an average of \$1032.70 (SD = \$2274.0) under the stated assumptions. In the cost-centered approach, we categorized patients into 4 quartile-based subgroups based on spending: high spenders (75% and up), medium spenders (50%-75%), low spenders (25%-50%), and zero spenders (0%-25%). The clinically focused approach detected 3 subgroups using the CP-learning algorithm based on clinical conditions and needs: high complexity, medium complexity, and low complexity. The **Table** displays the descriptive statistics of each subgroup, under both clinically focused and cost-centered approaches.

The Table reveals that the 2 medium subgroups—medium spenders and medium complexity—obtained from the clinical and cost approaches are in fact clinically quite similar. For example, the subgroups had, on average, 6.2 and 5.4 visits, and 5.6 and 5.3 unique diagnoses, but differed in the number of unique drugs (9.7 vs 3.6, respectively); and as an outcome measure, 16.7% and 14.1%, respectively, of the patients progressed beyond CKD stage 3. Even the average cost, \$274.6 and \$272.5, respectively, differed by just

FIGURE 1. Number of Patients by Spending, Color-Coded by Clinical Complexity

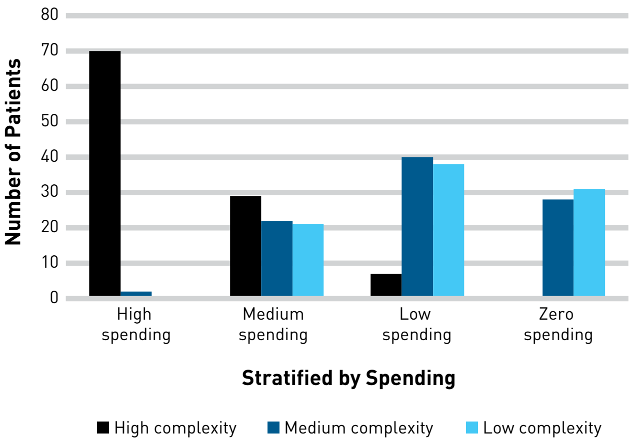
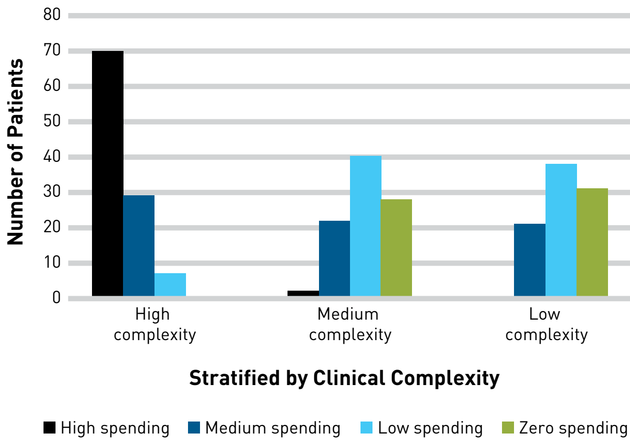
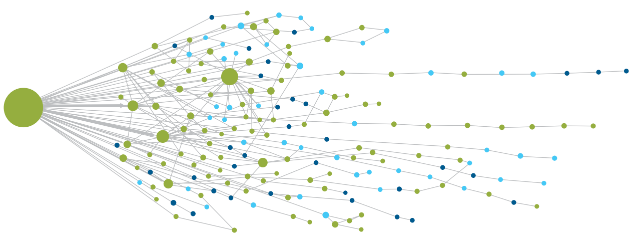


FIGURE 2. Variations in Cost Among Clinically Similar Patients

(a) Number of patients by clinical complexity, color-coded by spending



(b) Visualization of CPs in the medium complexity subgroup, color-coded by cost of co-pays of medications prescribed in visit



CP indicates clinical pathway.

over $2. However, what is strikingly different is the variation in costs between the 2 subgroups. Whereas the coefficient of variation of medium spenders was 0.73, that of the medium complexity group was 7.1—nearly a 10-fold increase.

Figure 1 displays the overlap in the assignment of patients to subgroups under clinically focused and cost-centered approaches. Although 97% of the high-spending patients were also part of the high-complexity group, 2 patients in the high-spending group—including the patient who spent the most, $18,801.40—were in fact assigned to the medium complexity group. One potential explanation, which needs to be verified by the clinicians, is that there might have been excessive spending of medical resources given these patients' clinical needs, which are at the medium-complexity level.

In contrast, **Figure 2** (parts a and b) displays the large variations in cost among clinically similar patients. Figure 2(b) displays patients' actual CPs generated for the medium-complexity subgroup. It can be pictured as a map of visits that patients experienced during the 2-year period. Each node in Figure 2 is a unique visit characterized by encounter type, diagnoses noted, and combinations of drug classes prescribed to patients. Starting with the dense, overlapped areas at the left-hand side of Figure 2(b), with all patients diagnosed with CKD stage 3, diabetes, and hypertension, but no other complications, the fanning pathways show the divergence from the common starting point as patients' disease progresses and complications emerge in diverse ways, for which varying treatments are provided. The size of the nodes and thickness of the edges reflect frequency of visits and transitions in the data. A larger node suggests that this is a common visit that many patients experience, and a thicker edge is an indication that the transition of visits, and accompanying change in clinical conditions or medication prescriptions, are observed among many patients. The color of each visit represents the spending category in quartiles: high (black), medium (dark blue), low (light blue), and zero (green).

DISCUSSION

There is growing recognition of the need for more precise risk-adjustment strategies, incorporation of evidence-based treatment variability, and increased use of data and information technology to facilitate patient engagement and shared decision making in promoting value-based payment systems.[4,16,17] In this preliminary study, we aimed to provide a generalizable framework to estimate the costs associated with actual care delivery, and to expose variations in care and cost. In particular, the results from this study showed significant variation in costs among patients who are clinically similar. A deeper analysis of these pathways may uncover the patterns and causes for these variations within subgroups to allow appropriate incorporation of this evidence into the development of future payment models and care delivery practices.

In this study, we analyzed data on CKD patients whose complex, chronic condition is an example of the high-need, high-cost care delivery context that is a continuing challenge to the healthcare system.[18] These patients require coordinated care due to their MCCs, and are a key population to be considered in policy design and implementation.[4] Our proposed framework specifically targets this patient population by modeling the co-progression of multiple clinical factors, treatments, and medications. We envision that an information technology–enabled tool based on the demonstrated methodology, once developed, deployed, and rigorously evaluated, can be used at the point of care by clinicians and patients to discuss available courses of treatment options, consider their potential efficacy projected at the cohort and personal level, and, equally important, build awareness of the costs associated with the entire course of treatment. Such tools may also provide policy makers and other stakeholders at healthcare practices access to data-driven evidence for innovative cost analyses.

Limitations

Although the methodology is generalizable to other health conditions and includes many clinical factors, a major limitation of our analysis is the accuracy and availability of data, particularly relevant to our cost-estimation approach. For illustrative purposes, we assigned all patients to a single Medicare plan and manually obtained the estimated co-pays from a website that provides prescription price information. Therefore, some of the drugs selected and prescribed to patients that were found to be excessively expensive may well be due to the potential discrepancy between our assumed insurance plan and patients' actual plans.

In addition, variations in CPs observed within each subgroup may be explained by unobserved variables, such as social and behavioral factors, as well as the more extensive health information that was not included in the current analysis. For example, the nature of our EHR dataset limits our ability to infer medication adherence; thus, the generated CPs assume perfect medication adherence, which we recognize to be unrealistic. There is also the possibility that the lack of adherence may lead to divergences in the CPs, but the current data fail to capture such associations. Furthermore, patients' conditions are gauged using *ICD-9-CM* codes recorded in the EHR; therefore, we did not distinguish patients by the severity of each condition. For instance, use of insulin, whose choice is a marker of severity in diabetes,[19] is often observed among the high-spending/complexity subgroup, but our CP-learning algorithm considered all patients with diabetes to have the same severity. Availability of such detailed, relevant data in the future—such as claims data and lab results—will help to overcome these limitations.

CONCLUSIONS

This preliminary analysis shows that patient subgroups generated by the CP-learning methods may be able to expose variations in costs among patients who are clinically similar, and vice versa, thereby facilitating future research to develop improved treatment plans within innovative payment models. ■

Acknowledgments

The authors are grateful to the physicians and staff of the community nephrology practice, who generously provided data from their electronic health record for this study. They particularly thank Pradip Teredesai, MD; Nirav Patel, MD; Qizhi Xie, MD, PhD; and staff members Linda Smith and Audra Barletta, who also gave important clinical and technical information about the data and the key characteristics of chronic kidney disease and its treatments.

Author Affiliations: Division of Health Informatics, Department of Healthcare Policy & Research, Weill Cornell Medicine of Cornell University (YZ), New York, NY; The H. John Heinz III College, Carnegie Mellon University (RP), Pittsburgh, PA.

Source of Funding: None.

Author Disclosures: The authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (YZ, RP); acquisition of data (YZ, RP); analysis and interpretation of data (YZ, RP); drafting of the manuscript (YZ); critical revision of the manuscript for important intellectual content (YZ, RP); statistical analysis (YZ, RP).

Address Correspondence to: Yiye Zhang, PhD, Assistant Professor, Division of Health Informatics, Department of Healthcare Policy & Research, Weill Cornell Medicine of Cornell University, 425 E 61st St, Ste 301, New York, NY 10065. E-mail: yiz2014@med.cornell.edu.

REFERENCES

1. Rein A. Navigating health care: why it's so hard and what can be done to make it easier for the average consumer. Changes in Health Care Financing and Organization website. http://www.hcfo.org/files/hcfo/HCFO-NavigatingHealthCare.pdf. Published 2007. Accessed November 1, 2016.
2. Alston C, Paget L, Halvorson G, et al. Communicating with patients on health care evidence. https://nam.edu/wp-content/uploads/2015/06/VSRT-Evidence.pdf. Published September 2012. Accessed November 1, 2016.
3. Jolly SE, Navaneethan SD, Schold JD, et al. Development of a chronic kidney disease patient navigator program. *BMC Nephrol.* 2015;16:69. doi: 10.1186/s12882-015-0060-2.
4. Blumenthal D, Abrams MK. Tailoring complex care management for high-need, high-cost patients. *JAMA.* 2016;316(16):1657-1658. doi: 10.1001/jama.2016.12388.
5. O'Connor AM, Wennberg JE, Legare F, et al. Toward the 'tipping point': decision aids and informed patient choice. *Health Aff (Millwood).* 2007;26(3):716-725.
6. O'Connor AM, Bennett CL, Stacey D, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev.* 2009(3):CD001431. doi: 10.1002/14651858.CD001431.pub2.
7. Whaley C, Schneider Chafen J, Pinkard S, et al. Association between availability of health service prices and payments for these services. *JAMA.* 2014;312(16):1670-1676. doi: 10.1001/jama.2014.13373.
8. Riggs KR, DeCamp M. Providing price displays for physicians: which price is right? *JAMA.* 2014;312(16):1631-1632. doi: 10.1001/jama.2014.7105.
9. Epstein AJ, Ketcham JD. Information technology and agency in physicians' prescribing decisions. *Rand J Econ.* 2014;45(2):422-448. doi: 10.1111/1756-2171.12057.
10. Chawla A, Westrich K, Matter S, Kaltenboeck A, Dubois R. Care pathways in US healthcare settings: current successes and limitations, and future challenges. *Am J Manag Care.* 2016;22(1):53-62.
11. Zhang Y, Padman R, Wasserman L, Xie Q. On clinical pathway discovery from electronic health record data. *IEEE Intelligent Systems*. 2015;30(1):70-75. doi: 10.1109/MIS.2015.14.
12. Zhang Y, Padman R, Patel N. Paving the COWpath: learning and visualizing clinical pathways from electronic health record data. *J Biomed Inform.* 2015;58:186-197. doi: 10.1016/j.jbi.2015.09.009.
13. Zhang Y, Padman R. Innovations in chronic care delivery using data-driven clinical pathways. *Am J Manag Care.* 2015;21(12):e661-e668.
14. Collins AJ, Foley RN, Gilbertson DT, Chen SC. United States Renal Data System public health surveillance of chronic kidney disease and end-stage renal disease. *Kidney Int Suppl (2011).* 2015;5(1):2-7.
15. Kovesdy CP, Kalantar-Zadeh K. Observational studies versus randomized controlled trials: avenues to causal inference in nephrology. *Adv Chronic Kidney Dis*. 2012;19(1):11-18. doi: 10.1053/j.ackd.2011.09.004.
16. Witkowski M, Higgins L, Warner J, Sherman M, Kaplan RS. How to design a bundled payment around value. *Harvard Business Review* website. https://hbr.org/2013/10/how-to-design-a-bundled-payment-around-value. Published October 3, 2013. Accessed November 1, 2016.
17. Mechanic R. Medicare's bundled payment initiatives: considerations for providers. American Hospital Association website. http://www.aha.org/content/16/issbrief-bundledpmt.pdf. Published January 19, 2016. Accessed November 1, 2016.
18. McCarthy D, Ryan J, Klein S. Models of care for high-need, high-cost patients: an evidence synthesis. *Issue Brief (Commonw Fund).* 2015;31:1-19.
19. Swinnen SG, Hoekstra JB, DeVries JH. Insulin therapy for type 2 diabetes. *Diabetes Care.* 2009;32(suppl 2):S253-S259. doi: 10.2337/dc09-S318.