Healthcare 7 (2019) 100366

Contents lists available at ScienceDirect

Healthcare

journal homepage: www.elsevier.com/locate/healthcare

Opinion paper

A delicate balance: Accountability for very high-cost patients in new payment models

Samyukta Mullangi[a], Matthew J. Press[b], Andrew M. Ryan[c,*]

[a] Department of Health Policy and Research, Weill Cornell Medicine, New York, NY, USA
[b] University of Pennsylvania Health System and Perelman School of Medicine, Philadelphia, PA, USA
[c] Department of Health Management & Policy University of Michigan School of Public Health, M3124 SPH II, 1415 Washington Heights, Ann Arbor, MI, 48109-2029, USA

ARTICLE INFO

Keywords:
Healthcare costs
Health services research
Health policy

ABSTRACT

Designing alternative payment models (APMs) to encourage providers to manage the needs of very complex patients while limiting providers' financial risk is a major challenge. Most APMs limit providers' financial risk for complex patients through an annual per-patient expenditure cap, above which medical expenses are excluded from financial performance calculations. We consider three strategies to better strengthen the balance between motivation and risk mitigation. First, risk adjustment could be improved. Second, expenditure caps can be tailored to very high-cost patients. Third, payers could allow providers to determine their own risk preferences through alternative reinsurance policies.

The costliest 5% of patients, who account for nearly half of medical spending, are a vulnerable group that need attention from provider groups.[1] High-cost patients often have multiple medical conditions, mental illness, or are at the end of life.[2] While some patients move in and out of this category as their health waxes and wanes, new data demonstrate the persistence of spending patterns: of those in the top 10% of spending, one third remain in the top 10% 5 years later.[3] High-cost beneficiaries also drive variation in spending within and across regions. Using spending data from the Medicare Beneficiary Annual Summary File, we calculated that spending for the median beneficiary in 2014 differed by only $506 between high- and low-spending regions (i.e., 75th vs. 25th percentile of Hospital Referral Regions (HRRs)). This difference increased to $9,848 for beneficiaries at the 95th percentile and $21,384 for beneficiaries at the 99th percentile (Fig. 1, Panel A). These spending patterns are persistent: spending at HRRs' 95th percentile between 2013 and 2014 is correlated at 0.99 (Fig. 1, Panel B). This suggests that care for high-cost patients reflects provider treatment patterns rather than just idiosyncratic health shocks for patients.

Current alternative payment models (APMs), such as accountable care organizations (ACOs) and bundled payments, often limit provider accountability for high-cost patients. Perhaps for this reason, ACOs have not disproportionately reduced spending for patients with greater clinical risk and expected expenditures.[4,5] When creating APMs, payers must shield providers from major financial losses incurred from the care of very high-cost patients, much of which may not be preventable. Payers must also be sure the incentives they create do not stifle the use of new and innovative technologies that may be clinically beneficial but expensive. However, providers must be rewarded for better managing these patients whose care is complex but frequently not well-coordinated and patient-centered.

The primary approach for extremely high-cost patients used by Medicare and other payers in APMs today is to define an annual per-patient expenditure level, either as a flat dollar amount or as a percentile, above which any medical expense is fully or partially excluded from the financial performance calculation. For example, in the Medicare Shared Savings Program, spending is truncated at the 99th percentile of the national fee-for-service population. Under Medicare's Comprehensive Care for Joint Replacement Model, a bundled payment program, spending that is two deviations greater than the regional spending target is capped. A similar strategy is employed under the Oncology Care Model (OCM), Medicare's bundled payment program for cancer care, which, when setting benchmark episode prices and calculating actual performance period episode expenditures, truncates expenditures at the 5th and 95th percentiles of the cancer-specific distribution of national episode expenditures.[6] These APMs also use additional stop-loss measures to cap total payments due from hospitals and physicians back to Medicare when total spending exceeds a spending threshold.

These checks are in place to protect provider organizations from severe financial losses, thereby encouraging participation in APMs, and

* Corresponding author.
E-mail address: amryan@umich.edu (A.M. Ryan).

https://doi.org/10.1016/j.hjdsi.2019.100366
Received 20 December 2018; Received in revised form 22 June 2019; Accepted 9 July 2019
Available online 24 September 2019

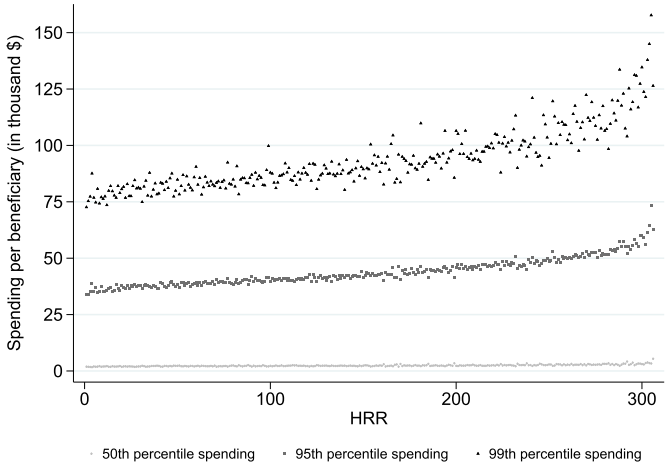

Fig. 1. Medicare spending per beneficiary across hospital referral regions.
Panel A. Medicare spending per beneficiary at the 50th, 95th, and 99th percentile by Hospital Referral Region, 2014
Notes: Data include observations from 30,848,623 Medicare beneficiaries from the Cost and Utilization Segment of the Master Beneficiary Summary File in 2014. Beneficiaries without 12 months of enrollment in Part A and Part B of traditional Medicare were excluded.
Panel B. Hospital Referral Region correlation between Medicare spending per beneficiary at 95th percentile in 2013 and 2014
Notes: Data include observations from 30,930,226 Medicare beneficiaries from the Cost and Utilization Segment of the Master Beneficiary Summary File in 2013 and 30,848,623 beneficiaries in 2014. The Pearson correlation coefficient is 0.99. Beneficiaries without 12 months of enrollment in Part A and Part B of traditional Medicare were excluded.

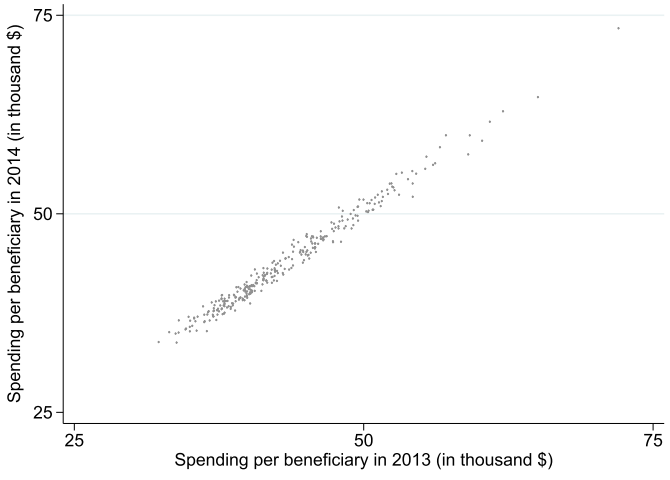

Fig. 1. (*continued*)

to help mitigate unintended consequences of APMs such as stinting or avoiding high-risk patients. But could retaining some modest performance risk for high-cost patients actually drive improvements in their care? While there are challenges in transmitting performance risk for these patients, we recommend three approaches that would reach a more optimal set of benefits and trade-offs.

First, risk-adjustment models should be improved to ensure that providers serving high-risk populations are not unfairly penalized within the financial performance calculation. While CMS has historically used coded diagnoses in a base year to reflect a population's cost pattern going forward, a more fluid methodology that incorporates specific trends in utilization may be more accurate. Applying machine learning to make prospective risk adjustment (predictions about severity of illness based on treatment patterns) can pare down human engineering in coding severity levels, which will result in improved accuracy.[7,8] Enhancing the data used for risk adjustment, particularly in hospitals, is also key. Compared to standard comorbidity coding, clinical and physiological data from Veterans Administration hospitals has been shown to vastly improve the quality of risk prediction.[9] In addition, oncology practices in OCM are providing clinical data to CMS through a registry that could be used in the future for more fine-tuned risk adjustment. Adjusting for social risk will also improve risk adjustment.[10]

Second, capping can be tailored to very high-cost patients. For example, the thresholds used for capping could vary across regions, diagnoses, or both. As we have shown, the 95th percentile of spending varies considerably across regions.[1] Adjusting the threshold for capping across regions, as is already done in the Comprehensive Care for Joint Replacement Model, would provide stronger incentives for providers in higher spending regions to manage high-cost patients. Similarly, patients' spending distribution varies across diagnoses. The OCM accounts for this by varying the spending thresholds by type of cancer, and adjusting this threshold across regions.[6] These strategies could also be employed for specific therapies. For instance, again in the OCM, episode prices are adjusted to account for the higher cost of certain new drugs when uptake of those drugs is higher in the OCM practice compared with non-OCM practices, but only at 80% of the difference. OCM practices are financially protected in their use of innovative (and expensive) new therapies but still have some "skin in the game" to use those therapies appropriately.

Finally, instead of its one-size-fits-all capping and stop-loss policies, CMS could encourage providers to accept more risk for high-cost patients. For instance, in exchange for higher shared savings, CMS could make providers responsible for some share of additional spending (e.g. 20%) among patients exceeding the spending threshold. In essence, instead of granting a single reinsurance policy for high-cost patients to providers, providers could be given a menu of options. Similar reinsurance policies were offered to provider groups facing capitation risk in the managed care era.[11] Such a policy could allow providers to determine their risk tolerance while guarding against catastrophic losses within APMs.

Providing stronger incentives to manage the care of high-cost patients must, of course, be balanced against the unintended consequence of providers seeking to avoid having these patients attributed to them in APMs. Improvements in risk adjustment, including adjustments for risk trajectories along with risk levels, would help to reduce the incentives to avoid patients. Prospective and multi-year attribution in ACOs would also help. Careful monitoring is required by payers to make sure that sure that providers are not caring for high-cost patients in response to incentives from APMs.

Very high-cost patients pose a vexing challenge to the potential for APMs to drive high-value care. These patients warrant special attention that is commensurate with their impact on healthcare spending. As payers shift risk to providers, we recommend enhancing risk adjustment, capping (or corridors), and reinsurance. Incentives, no matter how they are structured, have consequences. It is in everyone's interest to find that delicate balance between providers' accountability and financial protections.

Conflicts of interest

None declared for all co-authors.

Additional information

Dr. Press was previously employed by the Centers for Medicare and Medicaid Innovation.

Funding

None.

Prior presentations

None.

Acknowledgements

Contributors: We would like to thank Allan S. Detsky, MD PhD for his critical review of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.hjdsi.2019.100366.

References

1. Mitchell E, Machlin S. *Concentration of Health Expenditures and Selected Characteristics of High Spenders.* Rockville, MD: U.S. Civilian Noninstitutionalized Population; 2015 2015. Statistical Brief: #506 https://meps.ahrq.gov/data_files/publications/st506/stat506.shtml.
2. Wammes JJG, van der Wees PJ, Tanke MAC, Westert GP, Jeurissen PPT. Systematic review of high-cost patients' characteristics and healthcare utilisation. *BMJ Open.* 2018. https://doi.org/10.1136/bmjopen-2018-023113.
3. Hirth RA, Gibson TB, Levy HG, Smith JA, Calónico S, Das A. New evidence on the persistence of health spending. *Med Care Res Rev.* 2015. https://doi.org/10.1177/1077558715572387.
4. McWilliams JM, Chernew ME, Landon BE. Medicare ACO program savings not tied to preventable hospitalizations or concentrated among high-risk patients. *Health Aff.* 2017;36(12):2085–2093. https://doi.org/10.1377/hlthaff.2017.0814.
5. Markovitz A, Mullangi S, JM H, Nuliyalu U, Ryan AM. ACOs and the 1%: changes in spending among high-cost patients following the Medicare shared savings program. *J Gen Intern Med.* 2019;34(7):1116–1118https://doi.org/10.1007/s11606-019-04963-2.
6. Kline RM, Muldoon LD, Schumacher HK, et al. Design challenges of an episode-based payment model in oncology: the Centers for Medicare & Medicaid services oncology care model. *J Oncol Pract.* 2017. https://doi.org/10.1200/JOP.2016.015834.
7. Kan HJ, Kharrazi H, Chang HY, Bodycombe D, Lemke K, Weiner JP. Exploring the use of machine learning for risk adjustment: a comparison of standard and penalized linear regression models in predicting health care costs in older adults. *PLoS One.* 2019. https://doi.org/10.1371/journal.pone.0213258.
8. Rose S. A machine learning framework for plan payment risk adjustment. *Health Serv Res.* 2016. https://doi.org/10.1111/1475-6773.12464.
9. Render ML, Kim HM, Welsh DE, et al. Automated intensive care unit risk adjustment: results from a National Veterans Affairs study. *Crit Care Med.* 2003. https://doi.org/10.1097/01.CCM.0000055372.08235.09.
10. Joynt Maddox KE. Financial incentives and vulnerable populations — will alternative payment models help or hurt? *N Engl J Med.* 2018;378(11):977–979. https://doi.org/10.1056/NEJMp1715455.
11. Simon CJ, Emmons DW. Physician earnings at risk: an examination of capitated contracts. *Health Aff.* 1997. https://doi.org/10.1377/hlthaff.16.3.120.