

Effects of assessing the productivity of faculty in academic medical centres: a systematic review

Elie A. Akl MD PhD, Joerg J. Meerpohl MD, Dany Raad MD, Giulia Piaggio PhD, Manlio Mattioni MD PhD, Marco G. Paggi MD PhD, Aymone Gurtner PhD, Stefano Mattarocci PhD, Rizwan Tahir MD, Paola Muti MD MSc, Holger J. Schünemann MD PhD

Competing interests: None declared.

This article has been peer reviewed.

Correspondence to:
Elie A. Akl,
elieakl@buffalo.edu

CMAJ 2012, DOI:10.1503/
cmaj.111123

ABSTRACT

Background: Many academic medical centres have introduced strategies to assess the productivity of faculty as part of compensation schemes. We conducted a systematic review of the effects of such strategies on faculty productivity.

Methods: We searched the MEDLINE, Healthstar, Embase and PsycInfo databases from their date of inception up to October 2011. We included studies that assessed academic productivity in clinical, research, teaching and administrative activities, as well as compensation, promotion processes and satisfaction.

Results: Of 531 full-text articles assessed for eligibility, we included 9 articles reporting on eight studies. The introduction of strategies for assessing academic productivity as part of compensation schemes resulted in increases in clinical productivity (in six of six studies) in terms of clinical revenue, the work component of relative-value units (these units are non-monetary standard units of measure used to indicate the value of services provided),

patient satisfaction and other departmentally used standards. Increases in research productivity were noted (in five of six studies) in terms of funding and publications. There was no change in teaching productivity (in two of five studies) in terms of educational output. Such strategies also resulted in increases in compensation at both individual and group levels (in three studies), with two studies reporting a change in distribution of compensation in favour of junior faculty. None of the studies assessed effects on administrative productivity or promotion processes. The overall quality of evidence was low.

Interpretation: Strategies introduced to assess productivity as part of a compensation scheme appeared to improve productivity in research activities and possibly improved clinical productivity, but they had no effect in the area of teaching. Compensation increased at both group and individual levels, particularly among junior faculty. Higher quality evidence about the benefits and harms of such assessment strategies is needed.

Academic productivity can be defined as a measurable output of a faculty member related to clinical, research, education or administrative activities. Achieving the best possible academic productivity is essential for academic medical centres to maintain or nurture recognition and good reputation.¹ Furthermore, clinical productivity is essential for the survival of academic departments given the economic realities in medicine.²

Strategies for productivity assessment help in identifying highly productive faculty, determining areas for faculty and departmental improvement,³ and implementing processes for promotion and tenure.⁴ When coupled with reward schemes, these strategies may improve productivity and compensation at both individual and departmental levels. In the long-term, they may

enhance the ability to recruit and retain high-quality faculty and achieve the academic mission of the department. However, these strategies may have some unintended effects such as using time dedicated to education to do more clinical work.³ In addition, they may be challenging to implement.³

We conducted a systematic review of the effects of strategies introduced in academic medical centres to assess faculty productivity, compensation, promotion processes and satisfaction.

Methods

We defined productivity as a measurable activity or a measurable output of an activity of a faculty member. The considered areas of activity were clinical, research, teaching and administration.

Literature search

We searched the following databases from their date of inception up to October 2011: MEDLINE, Healthstar, Embase and PsycInfo. The search strategies included no language or date restrictions (for details, see Appendix 1, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.111123/-/DC1). In addition, we screened the reference lists of included articles and relevant papers and used the “related article” feature in PubMed to identify additional studies. We contacted the authors of studies that described strategies but did not report on their effects, seeking any unpublished data. We also searched websites of academic institutions and contacted experts in the field as part of the scoping for this study.

Eligibility criteria

We included studies that evaluated faculty members in academic medical centres; assessed the effects of a productivity assessment strategy introduced as part of appointment, promotion, compensation or incentive schemes; and had either no active comparison or a comparison with the introduction of another productivity assessment strategy.

Our main outcome measure was the effect on productivity in at least one of four areas (clinical, research, teaching and administration) either at the individual (faculty) or group (department) level. Additional outcomes of interest were the effects on faculty compensation, promotion processes and satisfaction. We were also interested in other relevant academic, educational, clinical and financial outcomes as well as negative unintended consequences of the strategies.

We excluded reports if they described a productivity assessment strategy but not its effects on the outcomes of interest; if they described interventions not used as part of a productivity assessment strategy (e.g., pay for performance, evaluation of the quality of faculty teaching); and if the strategies were related to concepts such as benchmarking and faculty development. We did not include conference proceedings.

Data selection and abstraction

Pairs of two reviewers independently screened the titles and abstracts of identified records for potential eligibility. We retrieved the full texts of citations judged to be potentially eligible by at least one reviewer. The groups of two reviewers then independently screened the full texts for eligibility and abstracted data using pilot-tested and standardized forms. They resolved their disagreements by discussion and, if necessary, with help from an arbiter. We contacted authors to verify our abstracted data.

Statistical analysis

We calculated the agreement between the reviewers for full-text screening using the kappa statistic. The studies used different outcome measurements and expressed results in various statistical formats, which precluded the conduct of meta-analyses. Thus, we summarized the results in narrative and tabular formats. We graded the quality of evidence using the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach⁵ and assigned a grade of “high,” “moderate,” “low” or “very low.”

Results

Study selection

Figure 1 shows the selection of studies for our review. The nine eligible articles reported on eight studies that evaluated eight different productivity assessment strategies.^{3,6-13} Agreement between reviewers for study eligibility was high (κ value = 0.75). The authors of the eight studies verified our abstracted data. We learned that

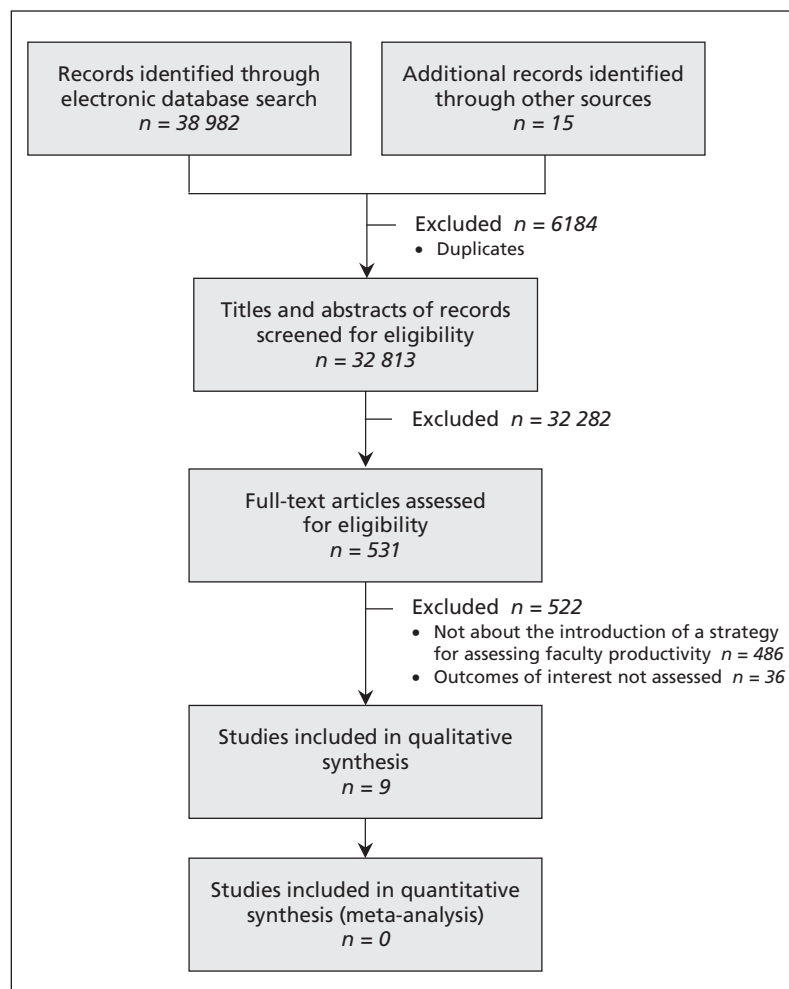


Figure 1: Selection of studies for the review.

seven of the strategies are still in use either in their original form or in a modified version.⁶⁻¹³

Appendix 2 (available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.111123/-/DC1) summarizes 35 strategies described in 36 reports that we excluded for not reporting on their effects. Authors of nine of these reports responded to our emails, indicating they had no data available on the effects of their strategies.

Features of the strategies

Details of the eight productivity assessment strategies evaluated in the included studies are summarized in Table 1. The strategies were reported between 1999 and 2008.

The development process was described for three strategies and included literature review, faculty input and creation of a special committee.^{3,6,11} The need for staff support for managing productivity assessment was reported for four strategies.^{3,6,8,11} Four strategies were described as offering either web-based access^{9,11} or regular feedback to faculty.^{8,10}

A variable number of the reports provided specific and reproducible descriptions of the strategies for assessing the four areas of productivity: clinical (eight studies^{3,6-11,13}), research (six^{3,6,9-11,13}), teaching (six^{3,6-8,11,13}) and administrative (four^{3,9,11,13}). The strategies used one or more methods for assessing the different areas of productivity, but the methods were not consistent across these strategies. The main measurement methods used were the relative-value unit (six studies^{3,6-9,11}), a rating scale determined by departmentally based standards (one¹³) and billable hours (one¹⁰). (The relative-value unit is a nonmonetary standard unit of measure used to indicate the value of services provided; for a specific service, the unit is multiplied by a conversion factor to calculate the total dollar amount assigned to that service.)

Features of the studies

Details of the included studies are summarized in Tables 2 and 3. All of the studies were conducted in the United States. In terms of design, all of the studies were conducted at a single institution and were designed for different institutional levels: six at the department level (four in primary care^{6,7-9} and two in anesthesiology^{10,11}), one at the school level (school of medicine³) and one at the level of academic centre (at the centre's schools of medicine, nursing and dentistry¹³). All of the studies compared data collected before with data collected after the implementation of the assessment strategy. Six studies explicitly reported a retrospective approach to data collection.^{3,6,7,10,11,13} The authors of the two remaining studies confirmed that data collection

was prospective.^{8,9} None of the studies included an active comparator.

The study populations were either restricted to primary care physicians⁶⁻⁹ or anesthesiologists,^{10,11} or they were more inclusive (i.e., one or more schools).^{3,13} Five studies reported the number of participants: 338 faculty in one study,⁹ and a range from 35 to 64 faculty in the others.^{6-8,10}

Of the eight studies, seven described the use of their productivity assessment as part of a compensation scheme.^{6-11,13} One of these seven studies reported the use of the assessment also as part of promotion and tenure standards.¹²

The numbers of studies assessing the outcomes of interest were as follows: clinical productivity (six^{3,6-9,11}), research productivity (six^{3,6,9-12}), teaching productivity (five^{6,8-11}), administrative productivity (none), compensation (five⁷⁻¹¹), promotion process (none) and other outcomes (two^{3,7}).

Risk of bias and overall quality of evidence

All eight studies^{3,13} used historical controls, and six collected data retrospectively (Table 2).^{3,6,7,10,11,13} None of the studies used validated outcome measures. Many of the studies reported imbalanced co-interventions (e.g., training in billing,^{7,8,10} introduction of a new rank system⁹ or improvement in research infrastructure¹³ concomitant with the introduction of the intervention) or other confounding factors (e.g., better payer contract,⁷ concurrent increase in federal grant funding⁹). Although we could not formally assess it, publication bias was possible. Similarly, included studies did not report on all of the outcomes of interest, or they reported on one but not all relevant aspects of an outcome (e.g., they reported on clinical productivity at the individual level but not at the group level). This may, at least in part, have led to selective outcome reporting. Consequently, we judged the overall quality of the evidence to be low.

Effects of the strategies

The effects of the strategies are summarized in Table 3.

Clinical productivity

All six studies assessing clinical productivity reported an increase in productivity in terms of clinical revenue, the work component of relative-value units, patient satisfaction and other departmentally used standards.^{3,6-9,11} Three studies reported increased productivity at the group level,^{3,7,8} one of these studies reported a simultaneous increase in productivity at the individual level,⁸ one reported a decrease at the individual level,³ and one did not report on individual productivity.⁷ Two studies reported decreased group

Table 1 (part 1 of 2): Description of strategies for the assessment of productivity of faculty reported by included studies

Study	Development of strategy	Areas of productivity assessed			
		Clinical	Research	Teaching	Administrative
Garson et al., 1999 ³	Faculty survey, metrics committee, faculty feedback and testing	<ul style="list-style-type: none"> Private patient care measured in FTEs and RVUs Community patient care Patient satisfaction 	<ul style="list-style-type: none"> Funded research Unfunded research Impact factor 	<ul style="list-style-type: none"> Total direct-contact teaching hours Teaching evaluation by learners 	Service on state/national committees
Cramer et al., 2000 ⁶	Comprehensive literature search, consensus process with faculty	<ul style="list-style-type: none"> RVUs based on clinical charges of faculty member's clinical care sessions and clinic precepting sessions RVUs for different services based on charge relative to that of CPT code 99123 (considered the unit RVU) 	<ul style="list-style-type: none"> RVU based on number of publications, grants, presentations For publications, RVU was based on the estimate of 240 hours to generate a paper worthy of publication (2 points/h x 240 h = 480 points) 	RVU based on number and type of teaching activity, course directorship, curriculum design, teaching award	No specific assessment method reported; compensation based on exceptional effort and special projects rewarded by the chair from an annual discretionary managerial incentive pool (5% of budgeted incentive funds)
Sussman et al., 2001 ⁷	Not reported	wRVU based on Medicare/Medicaid coding for evaluation and management	Not reported	<ul style="list-style-type: none"> Measured as "devoted time" Additional RVUs assigned per session of teaching 	Discretionary payment to primary care physicians (1% of productivity-derived salary) based on commitment to group activities (e.g., arranging CME for the group, leading medical management meetings)
Andreae et al., 2002 ⁸	Not reported	<ul style="list-style-type: none"> wRVU based on MGMA Physician Compensation and Productivity Survey¹⁴ Minimum wRVU = 70% (median wRVU for productivity of general pediatricians in MGMA survey¹⁴) Excess wRVU multiplied by \$/wRVU conversion factor¹⁴ Minimum wRVU and \$/wRVU adjusted annually 	No specific assessment method reported; compensation for research effort negotiated with the chair or unit director	wRVU teaching credit based on reported hours spent supervising medical students	No specific assessment method reported; compensation for administrative effort negotiated with the chair or unit director
Tarquinio et al., 2003 ⁹	Not reported	<ul style="list-style-type: none"> Based on RVUs in excess of those required to cover base salary Dollar value of benchmark RVU (\$/RVU) = benchmark salary (based on survey of peer institutions) divided by benchmark RVUs (based on MGMA academic RVU data or other sources) 	<ul style="list-style-type: none"> Based on the value of grants/gifts relative to base salary or NIH salary cap and on being principal investigator on a research grant "Research bonus" could be earned by clinician-educators and physician-scientists 	Not reported	Additional RVU credits for important administrative responsibilities such as division compliance expert, course director or fellowship program director
Miller et al., 2005 ^{10*}	Consultation with many groups, including private practitioners	<ul style="list-style-type: none"> "Credit" defined as billable hours collected minus expected hours Billable hours based on actual delivery of anesthetic clinical care Extra billable time given for specific clinical tasks (e.g., night or weekend work) but not for staffing more than one clinical site at a time 	No specific assessment method reported; instead of financial rewards, educators receive additional nonclinical time	No specific assessment method reported; incentives based on teaching evaluations	Not reported

Table 1 (part 2 of 2): Description of strategies for the assessment of productivity of faculty reported by included studies

Study	Development of strategy	Areas of productivity assessed			
		Clinical	Research	Teaching	Administrative
Reich et al., 2008 ¹¹	Chairperson advised by two committees to create valuations for activities	<ul style="list-style-type: none"> Point value scheme using the number of American Society of Anesthesiologists RVUs (ASA units) billed, adjusted for concurrent care Call assignments 	Point valuation scheme and valuation by a departmental committee based on faculty-submitted quarterly reports of academic activities	Point valuation scheme and valuation by the chair based on semi-annual evaluation by residents and peers	Point valuation scheme and valuation by the chair based on administrative functions assumed
Schweitzer et al., 2007 ¹² and 2008 ¹³	Committee of faculty and administrators	<ul style="list-style-type: none"> Measured against departmentally based standards Rating from 0 (unsatisfactory) to 4 (outstanding) 	<ul style="list-style-type: none"> Measured by grants and publications Rating from 0 (unsatisfactory) to 4 (outstanding) 	<ul style="list-style-type: none"> Measured against departmentally based standards Rating from 0 (unsatisfactory) to 4 (outstanding) 	<ul style="list-style-type: none"> Measured against departmentally based standards Rating from 0 (unsatisfactory) to 4 (outstanding)

Note: CME = continuing medical education, CPT = Current Procedural Terminology, FTE = full-time equivalent, MGMA = Medical Group Management Association, NIH = National Institutes of Health, RVU = relative-value unit (a nonmonetary standard unit of measure used to indicate the value of services provided; for a specific service, the RVU is multiplied by a conversion factor [dollar amount by unit of RVU] to calculate the total dollar amount assigned to that service), wRVU = work component of the RVU¹⁴ (under the Centers for Medicare and Medicaid Services, the RVU has three components: work, practice expense and professional liability insurance).
*Strategy also described in Feiner et al.¹⁵

productivity and increased individual productivity,^{6,11} with one describing an increasing workload undertaken by a stable number of full-time equivalent faculty.¹¹ One study reported increased individual productivity but did not report on group productivity.⁹

Research productivity

Five of the six studies assessing this outcome reported increased research productivity in terms of applications for funding,³ acquisition of funding,^{3,9,10,13} number of publications¹³ or some other measure.⁶ The sixth study reported that the rate of publication did not change after implementation of the strategy.¹¹ Two studies evaluated research productivity based on a point-based system using either relative-value units⁶ or a scheme developed by a departmental committee.¹¹

Teaching productivity

Of the five studies that assessed this outcome, two reported that the “educational output” was stable¹¹ or that the number of student and resident sessions did not change.⁸ Two studies reported no change in student or resident evaluations,^{8,9} and one reported improvement in such evaluations.¹⁰ One study reported a 4% drop in a mean composite score for teaching.⁶ However, that same study suggested that faculty were more willing to participate in educational activities.

Compensation

One of the five studies that assessed this outcome reported an increase in compensation at the group level by 8%,⁸ another reported an increase at the

individual level by 40%,¹¹ and a third study reported a “slight” increase at both levels.⁷ Two studies reported a change in compensation distribution in favour of junior faculty^{10,11} (defined in one study as assistant professors¹⁰).

Other outcomes

One study reported improvement or no change in patient satisfaction.⁷ That study showed varying levels of satisfaction among faculty; satisfaction was associated with “understanding of the plan” and inversely associated with years of service.

Only one study provided some evidence about shifts in area of productivity: faculty who were not productive in their research work shifted to patient care.³ None of the eligible studies assessed effects on administrative productivity or promotion processes.

Interpretation

We found that strategies introduced to assess faculty productivity as part of a compensation scheme in academic medical centres appeared to improve productivity in research activities and possibly improved clinical productivity but had no effect in the area of teaching. Compensation increased at both the group and individual levels, particularly among junior faculty. None of the studies assessed effects on administrative productivity or promotion processes.

Despite the low level of quality of the evidence, the data suggest that the assessment of faculty productivity, when coupled with an appropriate compensation or incentive scheme,

Table 2 (part 1 of 2): Design of studies evaluating the effects of the productivity assessment strategies

Study (setting)	Design	Population	Intervention	Risk of bias
Garson et al. ³ (Baylor College of Medicine, Houston, Texas)	Retrospective before–after comparison	17 clinical science departments and 8 basic science departments; <i>n</i> (faculty) not reported	<ul style="list-style-type: none"> • “Metrics process”; goals were set for each department for each area assessed (e.g., 10% increase in NIH grant dollars). • Individual and departmental performance data collected, analyzed and fed back yearly • Department-based compensation plans were in place (according to author; no further details provided) 	<ul style="list-style-type: none"> • Collected data retrospectively • Did not use validated outcome measures
Cramer et al. ⁶ (Department of Family Medicine, State University of New York at Buffalo, Buffalo, New York)	Retrospective before–after comparison	<i>n</i> (faculty) = 38–49	<ul style="list-style-type: none"> • Relative-value–based incentive plan • Points totalled quarterly, used as the divisor for that quarter’s available cash • Cash value per point multiplied by provider’s point total 	<ul style="list-style-type: none"> • Collected data retrospectively • Did not use validated outcome measures
Sussman et al. ⁷ (Division of General Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts)	Retrospective before–after comparison	Employed academic primary care physicians; <i>n</i> (faculty) = 64	<ul style="list-style-type: none"> • Productivity-based salary and bonus measures • Annual salary adjustments based on the prior year’s clinical productivity • Bonus measures paid quarterly based on medical management, quality of care, teaching and seniority • Magnitude of total bonus payments is based on percentage of the wRVU productivity-derived salary, to a maximum bonus of 10% 	<ul style="list-style-type: none"> • Collected data retrospectively • Did not use validated outcome measures • Reported imbalanced co-interventions (training in billing) and another confounding factor (better payer contract)
Andreae et al. ⁸ (Division of General Pediatrics, University of Michigan Health, Ann Arbor, Michigan)	Before–after comparison (data collection probably prospective)	<i>n</i> (faculty) = 35	<ul style="list-style-type: none"> • Productivity-based faculty compensation program that included a base salary and incentive payment • Once minimum wRVUs were generated to cover the base salary, excess wRVUs were converted to cash value as an incentive payment • Agreements for protected time or extramural salary support unaltered • Base salary independent of academic rank or years of service • Wages exclusive of benefits used for compensation 	<ul style="list-style-type: none"> • Collected data prospectively • Did not use validated outcome measures • Reported imbalanced co-interventions (training in billing)
Tarquino et al. ⁹ (Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee)	Prospective before–after comparison	12 clinical divisions; <i>n</i> (faculty) = 338	<ul style="list-style-type: none"> • Performance-based compensation plan that included a base salary and productivity adjustments • Productivity adjustments were paid when productivity exceeded amount to generate base salary • Clinician-educators and physician-scientists given up front 20% and 80%, respectively, of clinical RVU benchmark • Base salary based on that of the year before the intervention was implemented • Benchmark salary based on survey of peer institutions • Cash value per RVU was specialty-specific but unrelated to academic rank or track 	<ul style="list-style-type: none"> • Collected data prospectively • Did not use validated outcome measures • Reported imbalanced co-interventions (introduction of new rank system) and another confounding factor (concurrent increase in federal grant funding)
Miller et al. ¹⁰ (Department of Anesthesia and Perioperative Care, University of California at San Francisco, San Francisco, California)	Retrospective before–after comparison	Faculty with surgical and obstetric anesthesia responsibilities at three hospitals; <i>n</i> (faculty) = 58	<ul style="list-style-type: none"> • Productivity-based compensation including a base salary and productivity adjustments • Expected hours arbitrarily predetermined as 7.5 h/d of assignment • No information provided on conversion of credit to cash value 	<ul style="list-style-type: none"> • Collected data retrospectively • Did not use validated outcome measures • Reported imbalanced co-interventions (training in billing)

Table 2 (part 2 of 2): Design of studies evaluating the effects of the productivity assessment strategies

Study (setting)	Design	Population	Intervention	Risk of bias
Reich et al. ¹¹ (Department of Anesthesiology, Mount Sinai School of Medicine, New York, New York)	Retrospective before–after comparison	<i>n</i> (faculty) not reported	<ul style="list-style-type: none"> • Mission-based productivity compensation model with a base salary (based on rank and experience) and a supplemental pay (based on points earned) • Cash value of points based on finances of the department and total number of points generated by faculty in the preceding quarter • Compensation model placed 70% of total compensation at risk 	<ul style="list-style-type: none"> • Collected data retrospectively • Did not use validated outcome measures
Schweitzer et al. ^{12,13} (Health Sciences Center schools of dentistry, nursing and medicine, University of Louisville, Louisville, Kentucky)	Retrospective before–after comparison	<i>n</i> (faculty) not reported	<ul style="list-style-type: none"> • Recruitment of endowed chairs and their teams • New promotion and tenure standards • Salary incentives linked to research productivity • Post-tenure review: faculty given an unsatisfactory review were reviewed again 2 years later; failed reviews triggered the creation of a faculty development plan • Improvements to the research infrastructure • The chair's salary comprised a base salary and a salary supplement based on performance measures 	<ul style="list-style-type: none"> • Collected data retrospectively • Did not use validated outcome measures • Reported imbalanced co-interventions (improvement in research infrastructure)

Note: NIH = National Institutes of Health, RVU = relative-value unit, wRVU = work component of the RVU¹⁴ (see Table 1 footnote for description of RVU and wRVU).

can lead to positive systemic cultural change and help to achieve a department's mission.¹ For example, one of the included studies reported that the strategy resulted in non-research faculty volunteering for more teaching and administrative responsibilities, which freed up research faculty to engage in scholarly pursuits.¹²

Assessment of productivity can have some negative unintended consequences. Faculty might assume that an item not being evaluated is less important.³ For example, not considering teaching productivity, or weighing it much lower than clinical productivity, may affect faculty who teach and hurt the mission of the department. The productivity assessment strategy and the compensation scheme should lead to proper relative compensation for the different areas. That approach would ensure their alignment with the mission of department and the goals of its chair.

Another unintended consequence may be driven by gaming the system, intentionally or not, to score high on the productivity assessment with disregard to quality (e.g., aiming to get good evaluations rather than provide good teaching). Clinical care and teaching are two areas vulnerable to such effects. Appropriate and validated measurement tools as well as an adequate culture are needed to help avoid these situations.

There are a number of challenges to assessing productivity (Box 1). The primary challenge is how to measure productivity. Indeed, clinical

activities might not be equally intense (e.g., on-call time intensity might vary depending on the site), which makes availability a suboptimal measure.³ Similarly, reimbursement for clinical work may not be ideal because it can vary depending on the site (county v. private clinic). Also, it may be difficult to determine the equivalent values across areas of productivity (e.g., clinical v. research). The most common approach to achieving a fair assessment has been the use of relative-value units. Different approaches have been used to develop these units. Hilton and colleagues developed a base unit for each area of productivity (clinical, teaching, research and administration) and adjusted them so that the total points for the presumed productivity of a "top producer" in each of the areas would be equal.¹⁶ Cramer and colleagues proceeded by agreeing on a relative-value unit for clinical productivity (using a specific billing level) and then determined equal points for equal effort in the other areas by committee consensus.⁶ Others developed scales that focused on one area only, such as teaching.¹⁷

Although strategies for assessing productivity can be used solely as evaluation tools (for both departments and faculty), they are most often used along with productivity-based compensation programs to reward and stimulate productivity. Typically, these programs place a percentage of income at risk (e.g., 20%)¹¹ and reward productiv-

Table 3 (part 1 of 2): Effects of strategies on faculty productivity and compensation

Study	Clinical productivity	Research productivity	Teaching productivity	Compensation
Garson et al. ³	<ul style="list-style-type: none"> Total revenue increased 1% Revenue per patient-care FTE decreased 2% (patient-care FTE increased, according to author) 	<ul style="list-style-type: none"> Unfunded research decreased 12% Grant dollars per square foot increased 12% Percentage of basic researchers who held or applied for an NIH grant as principal investigator increased 35% Percentage of researchers who supported at least half of the portion of their salaries allocated to research increased 6% Faculty who had previously claimed research time with no research productivity shifted to patient-care time 	Not assessed	Not assessed
Cramer et al. ⁶	<ul style="list-style-type: none"> Mean clinical points dropped from 3008 in 1997 to 2547 in 1999 Clinical points per session rose from 9.8 in 1997 to 10.4 in 1999 	<ul style="list-style-type: none"> Mean research points rose from 524 in 1997 to 775 in 1999 	<ul style="list-style-type: none"> Mean teaching points dropped from 1195 in 1997 to 1146 in 1999 Course coordinators had easier time getting faculty participation at the predoctoral and resident levels 	Not assessed
Sussman et al. ⁷	<ul style="list-style-type: none"> 20% increase in wRVU volume Growth in wRVUs was "slightly out of proportion" to volume of visits; possibly related to an active coding-education and coding-auditing program Improvement or no change in patient satisfaction after introduction of the system 	Not assessed	Not assessed	<ul style="list-style-type: none"> Physicians' salaries and the overall clinical revenue "increased slightly" under the new system Salary decreases for under-performance were limited to annual adjustments
Andreea et al. ⁸	<ul style="list-style-type: none"> 89% of faculty increased clinical productivity Twice as many faculty as in previous year reached median for wRVU production As a group, clinical productivity increased 20% in the first year after implementation; this increase "exceeded the market's growth trends for primary care physicians" 	Not assessed	<ul style="list-style-type: none"> No change in the number of student and resident sessions No difference in students' evaluation of their educational experience 	<ul style="list-style-type: none"> Compensation for the group increased by 8% Percentage of increase in clinical productivity was 2.5 times greater than increase in compensation for the group; authors explained that this was due in part to the relative underproduction for the level of compensation at baseline
Tarquino et al. ⁹	<ul style="list-style-type: none"> Mean number of RVUs per clinician-educator grew by a compound annual growth rate of 8.1% before and 11.3% after implementation; corresponding numbers for clinicians in the medical centre were -3.4% and 3.1% 	<ul style="list-style-type: none"> Amount of research funding per physician-scientist grew by a compound annual growth rate of 9.3% before and 22.5% after implementation Ranking of department according to NIH awards improved from 26 to 12 (there were concurrent increases in NIH annual budgets) 	<ul style="list-style-type: none"> No change in evaluations by students or house staff, which remained very high (author communication) Clinician-educators and physician-scientists reported spending less time teaching (50% and 43% respectively) and less time on academic services (67% and 49% respectively); no association with receiving research bonus 	<ul style="list-style-type: none"> 8.6% of improvement in collections was attributable to improved billing or better payer contracts

Table 3 (part 2 of 2): Effects of strategies on faculty productivity and compensation

Study	Clinical productivity	Research productivity	Teaching productivity	Compensation
Miller et al. ¹⁰	Not assessed	<ul style="list-style-type: none"> Federal and other extramural funding sources dramatically increased 	<ul style="list-style-type: none"> Evaluations by medical students were "the best at UCSF" Authors reported that the residency program "flourished" 	<ul style="list-style-type: none"> Mean difference in total compensation narrowed between ranks Variability in compensation increased in all ranks, especially at the assistant professor level Distribution of total compensation "dramatically" changed in favour of junior faculty
Reich et al. ¹¹	<ul style="list-style-type: none"> Measure of clinical productivity per full-time equivalent increased by 31% ($p < 0.0001$) Measure of clinical productivity per location decreased by 10% ($p = 0.046$) 	<ul style="list-style-type: none"> Rate of publication by faculty did not change after implementation of the strategy 	<ul style="list-style-type: none"> Educational output was stable before and after implementation 	<ul style="list-style-type: none"> Mean compensation increased by 40% (95% CI 29.0%–50.4%) Compensation for instructors and assistant professors increased more than compensation for associate professors and professors ($p < 0.001$) both before and after implementation Increased mean supplemental pay with relatively fewer faculty, and increasing workload undertaken by stable number of full-time equivalent faculty Implementation of compensation model coincided with introduction of electronic point-of-care charge capture and submission system
Schweitzer et al. ^{12,13}	Not assessed	<p>In the school of medicine:</p> <ul style="list-style-type: none"> NIH grants increased from 39 to 127 during the 10-year period Annual NIH funding increased by a factor of 7 Ranking according to NIH awards improved from 95 to 73 (of 124) Total annual extramural funding increased from \$20 million to \$90 million (comparable increases observed in nursing and dentistry schools) Number of full journal articles increased by 60%, with an increase in impact (as measured by standard impact metrics) 	Not assessed	Not assessed

Note: CI = confidence interval, FTE = full-time equivalent, NIH = National Institutes of Health, RVU = relative-value unit, wRVU = work component of the RVU¹⁴ (see Table 1 footnote for description of RVU and wRVU), UCSF = University of California at San Francisco.

Box 1: Challenges to assessing productivity of faculty

- The best way to measure productivity is not clear
- Faculty may have little or no control over their productivity because of factors such as clinic population, scheduling, staffing and nonclinical demands.⁶ Similarly, productivity in research activities may be affected by factors such as the availability of biostatistical and methodologic support.
- Implementation of a productivity assessment strategy may be hampered by the lack of timely and accurate billing data; dedicated staff may be required for the collection, verification and analysis of data.^{3,6,8}
- In many cases productivity data are self-reported and may require regular auditing.⁶
- Concerns about the fairness, accuracy and timeliness of the process may negatively affect faculty buy-in and need to be adequately addressed.⁶
- The assessment of clinical productivity typically does not account for teamwork and may affect it negatively. As a result, some have advocated the assessment of productivity at the group level.⁷

ity beyond the expected minimum (marginal productivity).¹⁸ Some have argued that an incentive system focused on specific goals is more effective than a productivity-based program.¹

A number of factors can affect the success of productivity-based programs. Rewards should not be based on thresholds that reward only exceptionally productive faculty.⁶ At the same time, they need to be incremental to avoid hitting a “magic number” above which faculty lose incentive for being more productive.⁶ Ideally, reward schemes need to be aligned with and adjusted regularly to local and national trends for compensation,⁸ as well as to increases or decreases in departmental finances. The rewards should be valuable enough (in monetary value) to have an impact.¹⁸

Strengths and limitations

The main strength of this study is our adherence to the high standards of conducting and reporting systematic reviews.¹⁹ In particular, we used a thorough process for selecting studies, abstracting data and verifying the results directly with the authors of the included studies.

The main limitation of our systematic review is inherent in the available evidence and the low methodologic quality of the studies and their poor reporting. Another limitation is the generalizability of the results. First, all of the results were from academic centres in the United States.

Given different academic environments in other countries, the findings may not apply elsewhere. Second, the studies focused on primary care and anesthesiology, and therefore the results may not be applicable to other specialties.

Conclusion

Strategies introduced to assess faculty productivity as part of a compensation scheme appeared to improve research productivity and possibly improved clinical productivity but had no effect in the area of teaching. Compensation increased at both group and individual levels, particularly among junior faculty.

Departments planning on introducing a strategy to assess productivity need to carefully consider the associated challenges, the uncertain benefits and the potential unintended effects. Appendix 2 provides further examples of such strategies. For academic medical centres to make informed decisions about the future use of productivity assessment strategies, they will need higher quality evidence about the benefits and harms of those strategies. Although cluster randomized trials are conceptually feasible, they may not be practical. Large controlled observational or before–after studies with careful handling of confounding (through matching and adjustment) could provide higher quality evidence. Also, a central repository of strategies, processes and measurement tools would be ideal to assist academic leaders in designing their own programs.

References

1. Abouleish AE. Productivity-based compensations versus incentive plans. *Anesth Analg* 2008;107:1765-7.
2. Ashar B, Levine R, Magaziner J, et al. An association between paying physician-teachers for their teaching efforts and an improved educational experience for learners. *J Gen Intern Med* 2007;22:1393-7.
3. Garson A Jr, Strifert KE, Beck JR, et al. The metrics process: Baylor's development of a “report card” for faculty and departments. *Acad Med* 1999;74:861-70.
4. Santilli SM. Current issues facing academic surgery departments: stakeholders' views. *Acad Med* 2008;83:66-73.
5. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction — GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383-94.
6. Cramer JS, Ramalingam S, Rosenthal TC, et al. Implementing a comprehensive relative-value-based incentive plan in an academic family medicine department. *Acad Med* 2000;75:1159-66.
7. Sussman AJ, Fairchild DG, Coblyn J, et al. Primary care compensation at an academic medical center: a model for the mixed-payer environment. *Acad Med* 2001;76:693-9.
8. Andreae MC, Freed GL. Using a productivity-based physician compensation program at an academic health center: a case study. *Acad Med* 2002;77:894-9.
9. Tarquinio GT, Dittus RS, Byrne DW, et al. Effects of performance-based compensation and faculty track on the clinical activity, research portfolio, and teaching mission of a large academic department of medicine. *Acad Med* 2003;78:690-701.
10. Miller RD, Cohen NH. The impact of productivity-based incentives on faculty salary-based compensation. *Anesth Analg* 2005;101:195-9.
11. Reich DL, Galati M, Krol M, et al. A mission-based productivity compensation model for an academic anesthesiology department. *Anesth Analg* 2008;107:1981-8.

12. Schweitzer L, Eells TD. Post-tenure review at the University of Louisville School of Medicine: a faculty development and revitalization tool. *Acad Med* 2007;82:713-7.
13. Schweitzer L, Sessler DI, Martin NC. The challenge for excellence at the University of Louisville: implementation and outcomes of research resource investments between 1996 and 2006. *Acad Med* 2008;83:560-7.
14. Medical Group Management Association. *Physician compensation and production survey report: 1998 report based on 1997 data*. Englewood (CO): The Association; 1998.
15. Feiner JR, Miller RD, Hickey RF. Productivity versus availability as a measure of faculty clinical responsibility. *Anesth Analg* 2001;93:313-8.
16. Hilton C, Fisher W, Lopez A, et al. A relative-value-based system for calculating faculty productivity in teaching, research, administration, and patient care. *Acad Med* 1997;72:787-93.
17. Bardes CL, Hayes JG. Are the teachers teaching? Measuring the educational activities of clinical faculty. *Acad Med* 1995;70:111-4.
18. Lubarsky DA. Incentivize everything, incentivize nothing [editorial]. *Anesth Analg* 2005;100:490-2.
19. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;339:b2535.

Affiliations: From the Departments of Medicine (Akl, Raad, Tahir) and Family Medicine (Akl), State University of New York at Buffalo, Buffalo, NY; the Departments of Clinical Epidemiology and Biostatistics (Akl, Schünemann), Medicine (Schünemann) and Oncology (Muti), McMaster University, Hamilton, Ont.; the Department of Development of Therapeutic Programs (Mattarocci, Paggi), the Experimental

Research Center (Mattioni), the Experimental Oncology Department (Gurtner, Piaggio) and the Scientific Directorate (Muti), Regina Elena National Cancer Institute, Rome, Italy; and the German Cochrane Centre, Institute of Biostatistics and Medical Informatics, and the Pediatric Hematology and Oncology Clinic, Center for Pediatrics and Adolescent Medicine, University Medical Center Freiburg (Meerpohl), Freiburg, Germany

Contributors: Elie Akl, Aymone Gurtner, Stefano Mattarocci, Manlio Mattioni, Joerg Meerpohl, Marco Paggi, Giulia Piaggio, Dany Raad and Rizwan Tahir acquired the data. Elie Akl and Holger Schünemann analyzed and interpreted the data. Elie Akl, Paola Muti and Holger Schünemann provided administrative, technical or material support and supervised the study. Paola Muti and Holger Schünemann obtained funding. Elie Akl drafted the manuscript. All of the authors critically revised the manuscript for important intellectual content and approved the final version submitted for publication.

Funding: This study was supported by internal funds from the Regina Elena National Cancer Institute in Rome. The institute had no role in the design and conduct of the study; the collection, management, analysis and interpretation of the data; or the preparation, review and approval of the manuscript.

Acknowledgements: The authors thank Ms. Neera Bhatnagar and Mr. Edward Leisner for their assistance in designing the search strategy. They also thank Ms. Ann Grifasi and Mr. Eric Zinnerstrom for their administrative assistance.