BMJ Open

BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (http://bmjopen.bmj.com).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open: first published as 10.1136/bmjopen-2020-037610 on 19 July 2020. Downloaded from http://bmjopen.bmj.com/ on April 9, 2024 by guest. Protected by copyright.

BMJ Open

Survey of physician attitudes to using multisource feedback for competence assessment in Alberta

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-037610
Article Type:	Original research
Date Submitted by the Author:	09-Feb-2020
Complete List of Authors:	Ashworth, Nigel; University of Alberta, Kain, Nicole; College of Physicians & Surgeons of Alberta, Continuing Competence; Univeristy of Alberta, Medicine Jess, Ed; The College of Physicians & Surgeons of Alberta, Mazurek, Karen; The College of Physicians & Surgeons of Alberta
Keywords:	EDUCATION & TRAINING (see Medical Education & Training), Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, MEDICAL EDUCATION & TRAINING

SCHOLARONE™
Manuscripts

BMJ Open: first published as 10.1136/bmjopen-2020-037610 on 19 July 2020. Downloaded from http://bmjopen.bmj.com/ on April 9, 2024 by guest. Protected by copyright.

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

For peer review only

Survey of physician attitudes to using multisource feedback for competence assessment in Alberta

Nigel Ashworth, MB.ChB., MSc.[1,2], Nicole Kain, PhD.[1,2], Ed Jess, BA.[2] and Karen Mazurek, MD.[2]

[1] Department of Medicine, University of Alberta, Edmonton; [2] College of Physicians & Surgeons of Alberta, Edmonton

No funding, conflicts or competing interests to declare from any author.

Corresponding author: ashworth@ualberta.ca

Word Count: 1982

Keywords: Physicians, MSF, Multisource, Feedback, Attitudes

For peer review only

Abstract

Background: The use of multisource feedback (MSF) for assessing physician performance is widespread and rapidly growing. Findings from early very small research studies using highly selected participants suggest high levels of satisfaction and support. However, after nearly two decades of experience using MSF to evaluate all physicians in Alberta we are skeptical of this.

Objectives: To determine physicians' actual opinions of MSF using the entire physician population of Alberta, Canada

Design: On-line survey

Setting: Alberta, Canada

Participants: All physicians with a full license to practice in Alberta in 2016 (>9000 physicians).

Interventions: All participants were asked to grade how well they thought MSF was at assessing various aspects of physician performance using a 10-point scale. There was also a text response field for written comments.

Outcomes: Mean responses to quantitative questions. Qualitative analysis of text responses.

We analyzed the data using SPSS23 and NVivo11 and built a multivariate model highlighting the predictors of high and low opinions of MSF.

Results: Survey response rate was high for physicians (25%). The mean rating for how successful MSF was at assessing a variety of dimensions, varied from a low of 5.03/10 for medical knowledge, to a high of 6.38/10 for professionalism and communication. Canadian trained MDs rated MSF significantly lower on every dimension by approximately 20% compared to non-Canadian trained MDs.

Conclusions: Physicians have much lower opinions about the ability of MSF to measure any dimension of their performance than what has been suggested in the literature. Canadian-trained MDs have a particularly low opinion of MSF for reasons that remain unclear. The results of this survey offer a

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

serious challenge to the effectiveness of a program that is designed to promote self-reflection and performance improvement.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Strengths and limitations of this study

- MSF is of considerable interest to physicians Worldwide and is being used as part of QI and QA assessments
- This is the largest study of its kind with over 2200 physician responses
- This was a mixed methods study that allowed for text responses to give in depth insight into the quantitative responses
- The findings are accurate to within 2.4% of the actual mean value 99 times out of 100
- Multivariate statistical methods

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

Despite the widespread and rapidly growing use of multisource feedback (MSF) for both formative and summative assessments of physician competence and performance, the actual views of physicians regarding MSF have not been well explored. This is surprising, given that a key assumption of MSF is that the 'weak' areas or opportunities for practice improvement identified will encourage physicians to make changes and improve in those areas. If the physician being assessed is not supportive of the process, then one can hardly claim that the individual will be likely to make changes. A qualitative study of consultant physicians in the Netherlands identified 'lack of openness and constructive feedback' as major barriers for the success of a MSF program (1). Furthermore, it is well-established in the business literature that MSF programs can cause significant negative effects with a decreased performance in one-third of all MSF programs (2-4). Additionally, significant anxiety and distress can be imposed upon individual physicians who are told they are in the 'bottom 10%' (5). Several small research projects of selected volunteer physicians who were used as participants to test new MSF programs or components of programs, found high rates of satisfaction of around 70% (6-9). However, these results can hardly be said to represent the views of the majority of physicians. Such contradictory findings in the literature have created an urgent need to better understand physicians' attitudes and opinions regarding MSF.

Since 1999, all physicians in Alberta were mandated to undergo a competence/performance assessment at least once every five years. The program, also known as the physician achievement review (PAR) program, comprised of a MSF assessment involving questionnaires from eight peer physicians, eight co-workers, twenty-five patients and a self-evaluation. The physician's aggregated scores were compared to reference data and presented in a report that ranked the physician in comparison to physicians with a similar practice on different aspects of performance (similar to CanMEDS dimensions of medical expert, communication, professionalism, and etcetera)(10). The terms 'multisource feedback (MSF)' and 'physician achievement review (PAR)' can essentially be used interchangeably, but for simplicity we use MSF exclusively in this paper.

For peer review only

We report on the findings of a survey of the entire physician membership in Alberta in 2016. The objectives of this survey were: 1) to discover physicians' actual opinions regarding MSF; 2) to see if different types of physicians (e.g. males v females, older versus younger) varied in their opinion of MSF.

Methods

This survey, as a component of a larger evaluation and re-design of the College of Physicians & Surgeons of Alberta (CPSA)'s continuing competence programs, received ethical approval from the University of Alberta joint Research Ethics Board. Participants gave informed consent at the time of the survey.

This was an electronic on-line survey sent to all physicians fully licensed to practice medicine in Alberta, Canada in 2016 (n=approximately 9,000). The survey was pilot tested several times and then sent out in two 'waves' over the course of 2-months.

The survey consisted of twelve questions asking the physician how successful MSF was at assessing different areas of physician competence (medical knowledge, clinical skills, communication skills, professionalism, and etcetera) in their opinion. There was also a text response field for physicians to provide written feedback. Finally, because the survey was anonymous, we asked respondents to provide information on their gender, year of graduation, whether they were a Family Medicine Physician/General Practitioner (FM), if they practiced solo, and whether they obtained their medical degree in Canada (Canadian MD). We used a 10-point modified Likert response scale to answer with three anchors (1= "not at all", 5 "fairly" and 10= "extremely").

Quantitative responses were analyzed using Statistical Package for the Social Sciences (SPSS) version 23. Descriptive, bivariate, and multivariate general linear model analysis was performed. Qualitative (open-ended) text responses were analyzed using NVivo11™ qualitative data analysis software.

We didn't involve patients or public in the design of this study.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Results

2215 physicians responded out of 9021 fully registered Alberta physicians (response rate of 24.6%). There is a 99% probability that the results obtained are within approximately 2.4% of the actual value for each of the survey questions.

Wave 1 consisted of 1387 physicians and wave 2 consisted of 828 physicians [Table 1]. The waves were significantly different from each other in terms of age and proportion of Canadian MDs with wave 1 being slightly longer in practice by a mean of four years, and having approximately 6% more Canadian MDs. Compared to the entire physician population, respondents (wave 1 and 2) were slightly longer in practice, more female, more FM, more Canadian MDs, and substantially less solo-practicing physicians [Table 1].

The mean rating for how successful MSF was at assessing a variety of dimensions, varied from a low of 5.03/10 for medical knowledge, to a high of 6.38/10 for professionalism and communication [Table 2]. A multivariate general linear model was created using the twelve survey questions as the dependent variables, and years in practice, Canadian MD, solo practice, family physician, and gender as the independent variables. The model showed that only Canadian MDs had significantly different opinions compared to non-Canadian-trained MDs [Table 3]. All of the other variables were not significant in the model and there were no significant interaction terms.

Qualitative content analysis of the 303 responses to the open-ended text box asking for "Additional Comments & Ideas" regarding MSF, revealed that respondents overwhelmingly regarded MSF as negative. Overarching themes included the opinions that MSF is:

1) A waste of time and/or resources ("To me the [MSF] is an incredible waste of time and money. I do not know one MD who actually pays it much attention, either filling it out or implementing changes");

2) Irrelevant or useless ("I believe that [MSF] is really useless as it currently operates. Major changes required.");

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

3) Subjective, biased ("The responses are biased to some or a large degree as the patients and peers that assess are chosen by the person being reviewed. If the person being reviewed is in an influential position, I'm not sure how accurate the review really is. I try to be as honest as I can when reviewing others, but wonder if that is always the case."); and

4) MSF can cause significant distress ("[I] find it[MSF] an extremely stressful exercise that I do every 5 years to maintain my license.").

Discussion

Surveys are notoriously difficult in physician groups who are often bombarded with multiple different survey requests on an almost daily basis (11,12). Previous surveys from CPSA have met with response rates of about 10% and this may not be unusual even for physician professional organizations(13). Our response rate of 25% was more than double the response rate of any previous survey by the CPSA, indicating that this is clearly an important topic for Alberta physicians. There is growing evidence that good quality and reliable information can be obtained from surveys with lower response rates (20-25%) and in fact they may be superior to higher response rate surveys(14-17). One major reason being that initial 'non-responders' who are 'forced' to answer do not do so in a benevolent way which can seriously undermine the validity of the findings. Secondly there was very little if any evidence to support that the typical minimum response rate should be 60-65%, this was anecdotal. Holbrook et all analyzed results from 81 national surveys (with response rates from 5-54%) and found the extra expense and energy invested in trying to boost response rates may not result in worthwhile improvements in reliability or validity(18).

Statistically, there was a difference between the different waves of respondents and between the entire population of Alberta physicians in terms of demographics, but given the large sample sizes, it may be very easy to find statistical differences even when the actual difference is small. Given that we know the composition of the responders and all Alberta physicians, we can predict the potential effects of any differences. Looking at the mean differences, wave 1 and 2 were very similar in composition between the entire population and the two waves, with one notable exception: there were appreciably less

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

physicians in solo practice who answered the survey (20% versus 30%). Given that the responses from solo physicians were identical to 'non-solo' physicians (with the exception that they rated 'group feedback' measures less importantly for obvious reasons), the validity of the overall results is established. Compared to the entire population, the respondents had practiced slightly longer, which had no effect on the survey responses. The respondent population was also composed of slightly more females, which might skew the questions on whether MSF 'inspires reflection' lower; slightly more Canadian-trained MDs, which might skew all opinions on MSF to be worse; and slightly more family physicians, which would skew most opinions on MSF to be slightly better. Thus, the overall effect of these differences is not expected to decrease the generalizability of the results to Alberta's physicians.

The CPSA research team was genuinely surprised to find such a low/mediocre rating for MSF (on average, 5-6/10), given the previous much more optimistic data from published research studies and non-independent small surveys. Results from this research indicate that substantial efforts are required to investigate and potentially improve the utility of MSF in Alberta physicians. The findings from previous research studies are greatly optimistic about physicians' acceptance and support of the MSF program (PAR from 1999 – 2015) in Alberta. Successful MSF programs in business tend to involve considerable investment in training reviewers beforehand, educating participants about the potential benefits of MSF, and providing feedback directly to the participant to help with understanding and facilitate actionable change. These strategies were not used in Alberta with the PAR program, and perhaps those who are considering applying MSF-type programs to physicians, should incorporate the valuable successes and variable lessons learned from the business milieu.

Furthermore, the substantially lower ratings by Canadian-trained MDs on almost every dimension of MSF were an even greater surprise to the research team. Again, considerable thought and investigation is required to assist in explaining why there is such a marked difference in opinion between these two groups. One could surmise that Canadian-trained MDs are more familiar with performance assessment (and MSF in particular), and can be much more critical in their responses. Conversely, non-Canadian-trained MDs might feel more grateful or appreciative for the opportunity of receiving feedback,

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

particularly if they had not been exposed to similar forms of feedback in their country of medical training. Additionally, non-Canadian-trained MDs might feel more vulnerable, or less comfortable criticizing the CPSA even though the study was anonymous and without consequence.

Given the way respondents rated MSF in Alberta, it is difficult to see how MSF could inspire the majority of physicians to make any significant changes to their practice. The mean score for "inspires reflection in practice" was only fair at 5.69/10. About 30% of physicians did rate the MSF 70% or higher, causing us to question how these physicians differ from those who rated MSF lower. The statistical analysis found that these physicians are more likely to be non-Canadian-trained MDs. However, the business literature would suggest that these "high rating" respondents may have unique personality types that are receptive to MSF (19). Unfortunately, because the survey was anonymous, the personalities of the "high rating" respondents cannot be further explored. Our findings complement research already completed in Alberta, where 72% of a cohort of surgeons taking part in the PAR program indicated initially that they were "contemplating or [had] initiated change on the basis of multisource feedback…"(20). However, after following-up with the cohort three months later, the same authors concluded that "surgeons made few changes in practice in response to feedback data"(21).

Physicians in Alberta have much lower opinions regarding the ability of MSF to measure the dimensions of performance than previously documented in the literature and internal reports. Canadian-trained MDs have a particularly low opinion of MSF for reasons that remain unclear. The results of this survey offer a serious challenge to the effectiveness of a program that is intended to promote self-reflection and performance improvement.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 1: Demographics of responders and non-responders

	Wave 1 (n=1387)	Wave 2 (n=828)	Wave1 versus Wave2 significance	All (n=9021)	Wave 1+2 versus All significance
Years in practice (sd)	26.0 (11.7)	22.9 (11.8)	<0.001	22.0 (12.2)	<0.001
Gender (% female)	38.8	41.4	NS (0.37)	37.2	0.02
Family physician (%)	51.8	57.1	NS (0.07)	48.5	0.001
Solo practice (%)	19.8	18.8	NS (0.67)	30.5	<0.001
Canadian MD (%)	73.8	67.3	0.02	66.6	<0.001

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2: Mean rating for question "Please rate how successful the existing [multisource feedback] program is in assessing the following dimensions. (1-10 where 1=not at all, 5=fairly, and 10= extremely)"

Dimension	Mean Score	SD
Medical knowledge	5.03	2.50
Clinical skills	5.17	2.47
Communication skills	6.38	2.42
Practice administration	5.36	2.47
Patient management	5.68	2.46
Adherence to standards of practice	5.34	2.52
Professionalism	6.38	2.47
Team functioning	6.15	2.54
Easy to participate	6.20	2.66
Inspires reflection in practice	5.69	2.63
Motivates clinical practice improve	5.58	2.62
Provide a learning opportunity	5.47	2.66
Overall	5.70	

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3: Mean rating for question "Please rate how successful the existing [multisource feedback] program is in assessing the following dimensions. (1-10 where 1=not at all, 5=fairly, and 10= extremely)" stratified by Canadian or Non-Canadian trained MD.

Dimension	Non-Canadian trained MD	Canadian trained MD	Multivariate analysis (p value)
Medical knowledge	5.71	4.74	<0.0001
Clinical skills	5.72	4.92	<0.0001
Communication skills	6.61	6.31	NS
Practice administration	6.04	5.08	<0.0001
Patient management	6.34	5.41	<0.0001
Adherence to standards of practice	6.11	4.98	<0.0001
Professionalism	6.73	6.30	NS
Team functioning	6.52	6.02	NS
Easy to participate	6.79	6.03	<0.0001
Inspires reflection in practice	6.35	5.43	<0.0001
Motivates clinical practice improve	6.29	5.31	<0.0001
Provide a learning opportunity	6.1	5.26	<0.0001
Overall	6.19	5.48	

NS = Not significant at 0.05 level

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Acknowledgements

Tanya Northfield for doing all the hard work and her endless patience with all of us. Dr. Ernie Schuster for the initial idea and help early on with the project.

Conflicts of interest

All the authors work for CPSA who ran the physician achievement review (PAR) program, a MSF based assessment program for 18 years in Alberta.

Author contributions:

All the authors were involved in the initial design of the study. Ashworth and Kain performed the analyses and initial write up. All the authors wrote and edited the finished manuscript.

Data availability statement: All data requests should be submitted to the corresponding author for consideration. Access to anonymized data will not be possible

Funding: None

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Bibliography

1. Overeem K, Wollersheim H, Driessen E, Lombarts K, van de Ven G, Grol R, et al. Doctors' perceptions of why 360-degree feedback does (not) work: a qualitative study. Med Educ. 2009 Sep;43(9):874–82.

2. Ilgen D, Davis C. Bearing Bad News: Reactions to Negative Performance Feedback. Applied Psychology. 2000;49(3):550–65.

3. Nowack KM, Mashihi S. Evidence-based answers to 15 questions about leveraging 360-degree feedback. Consulting Psychology Journal: Practice and Research. 2012;64(3):157–82.

4. Kluger AN, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin. 1996;119(2):254–84.

5. Atwater LE, Brett JF, Charles AC. Multisource feedback: Lessons learned and implications for practice. Hum Resour Manage. Wiley Subscription Services, Inc., A Wiley Company; 2007;46(2):285–307.

6. Sargeant JM, Mann KV, Ferrier SN, Langille DB, Muirhead PD, Hayes VM, et al. Responses of rural family physicians and their colleague and coworker raters to a multi-source feedback process: a pilot study. Acad Med. 2003 Oct;78(10 Suppl):S42–4.

7. Violato C, Marini A, Toews J, Lockyer J, Fidler H. Feasibility and psychometric properties of using peers, consulting physicians, co-workers, and patients to assess physicians. Acad Med. 1997 Oct;72(10 Suppl 1):S82–4.

8. Lockyer JM, Violato C, Fidler H. A multi source feedback program for anesthesiologists. Can J Anaesth. 2006 Jan;53(1):33–9.

9. Ramsey PG, Carline JD, Blank LL, Wenrich MD. Feasibility of hospital-based use of peer ratings to evaluate the performances of practicing physicians. Acad Med. 1996 Apr;71(4):364–70.

10. RCPSC. CanMEDS 2015 Physician Competency Framework. Frank J, Snell L, Sherbino J, editors. 2015 Oct pp. 1–17.

11. Tambor ES, Chase GA, Faden RR, Geller G, Hofman KJ, Holtzman NA. Improving response rates through incentive and follow-up: the effect on a survey of physicians' knowledge of genetics. Am J Public Health. American Public Health Association; 2011 Oct 7;83(11):1599–603.

12. Cunningham CT, Quan H, Hemmelgarn B, Noseworthy T, Beck CA, Dixon E, et al. Exploring physician specialist response rates to web-based surveys. BMC Med Res Methodol. 2nd ed. 2015 Apr 9;15(1):32.

13. CPSA. CPSA Survey 2014. Unpublished. 2014 May pp. 1–228.

14. Keeter S, Kennedy C, Dimock M, Public JBJO, 2006. Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. academicoupcom

.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

15. Visser PS, Krosnick JA, Opinion JMP, 1996. MAIL SURVEYS FOR ELECTION FORECASTING? AN EVALUATION OF THE COLUMBUS DISPATCH POLL. academicoupcom

.

16. Curtin R, Presser S, Singer E. The effects of response rate changes on the index of consumer sentiment. Public Opin Q. 2000;64(4):413–28.

17. Choung RS, Locke GR, Schleck CD, Ziegenfuss JY, Beebe TJ, Zinsmeister AR, et al. A low response rate does not necessarily indicate non-response bias in gastroenterology survey research: a population-based study. J Public Health. Springer-Verlag; 2012 Apr 25;21(1):87–95.

18. Holbrook A, Krosnick J, Pfent A. The Causes and Consequences of Response Rates in Surveys by the News Media and Government Contractor Survey Research Firms. In: Lepkowski JM, editor. Advances in telephone survey methodology [Internet]. New York: Wiley; 2007. pp. 1–92. Available from: https://pprg.stanford.edu/wp-content/uploads/2007-TSMII-chapter-proof.pdf

19. Smither JW, London M, Reilly RR. Does Performance Improve Following Multisource Feedback? a Theoretical Model, Meta-Analysis, and Review of Empirical Findings. Personnel Psychology. 2005 Mar;58(1):33–66.

20. Violato C, Lockyer J, Fidler H. Multisource feedback: a method of assessing surgical practice. BMJ. 2003 Mar 8;326(7388):546–8.

21. Lockyer J, Violato C, Fidler H. Likelihood of change: a study assessing surgeon use of multisource feedback data. Teach Learn Med. 2003;15(3):168–74.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in BMJ Open and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

For peer review only

BMJ Open

Survey of physician attitudes to using multisource feedback for competence assessment in Alberta

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-037610.R1
Article Type:	Original research
Date Submitted by the Author:	11-May-2020
Complete List of Authors:	Ashworth, Nigel; University of Alberta, Kain, Nicole; College of Physicians & Surgeons of Alberta, Continuing Competence; Univeristy of Alberta, Medicine Jess, Ed; The College of Physicians & Surgeons of Alberta, Mazurek, Karen; The College of Physicians & Surgeons of Alberta
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Qualitative research, Communication
Keywords:	EDUCATION & TRAINING (see Medical Education & Training), Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, MEDICAL EDUCATION & TRAINING

SCHOLARONE™
Manuscripts

BMJ Open: first published as 10.1136/bmjopen-2020-037610 on 19 July 2020. Downloaded from http://bmjopen.bmj.com/ on April 9, 2024 by guest. Protected by copyright.

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

For peer review only

Survey of physician attitudes to using multisource feedback for competence assessment in Alberta

Nigel Ashworth, MB.ChB., MSc.[1,2], Nicole Kain, PhD.[1,2], Ed Jess, BA.[2] and Karen Mazurek, MD.[2]

[1] Department of Medicine, University of Alberta, Edmonton; [2] College of Physicians and Surgeons of Alberta, Edmonton

No funding, conflicts or competing interests to declare from any author.

Corresponding author: ashworth@ualberta.ca

Word Count: 2914

Keywords: Physicians, MSF, Multisource, Feedback, Attitudes

For peer review only

Abstract

Background: The use of multisource feedback (MSF) for assessing physician performance is widespread and rapidly growing. Findings from early very small research studies using highly selected participants suggest high levels of satisfaction and support. However, after nearly two decades of experience using MSF to evaluate all physicians in Alberta we are skeptical of this.

Objectives: To determine physicians' actual opinions of MSF using the entire physician population of Alberta, Canada

Design: On-line survey

Setting: Alberta, Canada

Participants: All physicians with a full license to practice in Alberta in 2015.

Interventions: All participants were asked to grade how well they thought MSF was at assessing various aspects of physician performance using a 10-point Likert-type scale. There was also a text response field for written comments.

Outcomes: Mean responses to quantitative questions. Qualitative content and thematic analysis of open-ended text responses.

We analyzed the data using SPSS23 and NVivo11 and built a multivariate model highlighting the predictors of high and low opinions of MSF.

Results: Survey response rate was high for physicians with 2215 responses (25%). The mean rating for how successful MSF was at assessing a variety of dimensions, varied from a low of 5.03/10 for medical knowledge, to a high of 6.38/10 for professionalism and communication. Canadian trained MDs rated MSF significantly lower on every dimension by approximately 20% compared to non-Canadian trained MDs.

Conclusions: Alberta physicians have much lower opinions about the ability of MSF to measure any dimension of their performance than what has been suggested in the literature. Canadian-trained MDs

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

have a particularly low opinion of MSF for reasons that remain unclear. The results of this survey offer a serious challenge to the effectiveness of a program that is designed to promote self-reflection and performance improvement.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Strengths and limitations of this study

- MSF is of considerable interest to physicians Worldwide given it is increasingly being used as part of QI and QA assessments
- This is the largest study of its kind to date
- We employed a mixed methods analysis that allowed for open-ended text responses to give in-depth insight into the quantitative responses
- We used multivariate statistical methods to analyze the quantitative data
- Survey was anonymous but some physicians may still have been uncomfortable giving honest answers to the medical regulator

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

Multisource feedback (MSF) also known as 360-degree or multi-source or multi-rater feedback has been used as a tool for quality improvement and quality assurance in healthcare for over 20 years and in the business world for even longer.(1, 2) One advantage of MSF is that it can provide a physician with a much broader range of personal feedback (e.g. medical expert, communication, administration, professionalism) from many different sources (patient, co-worker, peer, self, leader) than might be obtained traditionally.(1, 3) In addition, because feedback from many patients and multiple peers and co-workers is typically obtained and aggregated, then this should average out any more extreme outlier views that assessors might express, hopefully giving a more balanced overall view.(4, 5) The overall psychometric properties, in terms of the reliability, validity and feasibility of the tool have been reported to be good.(3, 6, 7)

However, despite the widespread and rapidly growing use of MSF for both formative and summative assessments of physician competence and performance, the actual views of physicians regarding MSF have not been well explored. Previous studies of physician attitudes to MSF in Canada and Worldwide have shown very high rates of satisfaction with MSF, for example in one paper from Alberta 100% of 24 physicians felt that "[MSF] was a helpful educational exercise…".(3) Another study of 308 physician volunteers found 2/3 of those who responded to a post study satisfaction questionnaire (numbers not given) said they were contemplating or had made changes to their practice.(1) Other studies have typically surveyed similar small numbers of highly self-selected physicians.(8-10) A survey of 249 junior doctor attitudes in the UK found positive attitudes to MSF but low perceived opinion about the effectiveness.(11) This is surprising, given that a key assumption of MSF is that the 'weak' areas or opportunities for practice improvement identified will encourage physicians to make changes and improve in those areas. If the physician being assessed is not supportive of the process, then one can hardly claim that the individual will be likely to make changes. A qualitative study of consultant physicians in the Netherlands identified 'lack of openness and constructive feedback' as major barriers for the success of a MSF program.(12) Similarly, a qualitative study of opinions about MSF amongst

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

allied health staff and pediatric residents in Canada showed strong interest in the concept of MSF but significant potential barriers to success, such as poorly defined roles and responsibilities, perceptions of expertise, hostile hospital culture and negative interprofessionalism and power dynamics.(13). Only 1/16 studies looking at the effectiveness of MSF in physicians identified any significant positive change in actual behavior and in that study the treatment group also received "…a tailored coaching session to assist in identifying their strengths and weaknesses and in setting specific behavioural goals".(14-16) Furthermore, it is well-established in the business literature that MSF programs can cause significant negative effects with a decreased performance in up to one-third of all MSF programs due in part to an "exacerbation [of] bureaucracy, heightening [of] political tensions, and consumption [of] enormous numbers of hours".(17-20) Also giving good feedback is not easy, ideally the source should be credible and the information should be 'SMART': specific, measurable, attainable, relevant, and time-bound.(21) Additionally, significant anxiety and distress can be imposed upon individual physicians who are told they are in the 'bottom 10%.(22) Several small research projects of selected volunteer physicians who were used as participants to test new MSF programs or components of programs, found high rates of satisfaction of around 70%.(3, 9, 23-25) However, these results can hardly be said to represent the views of the majority of physicians. Such contradictory findings in the literature have created an urgent need to better understand physicians' attitudes and opinions regarding MSF.

Since 1999, all physicians in Alberta were mandated to undergo a competence/performance assessment at least once every five years, in accordance with Alberta's Health Professions Act.(26) The assessment, also known as the physician achievement review (PAR) program, comprised a MSF assessment involving questionnaires from eight peer physicians (colleagues), eight co-workers, (e.g. nurses, receptionists, physiotherapists), twenty-five patients and a self-evaluation.(27) The physician's aggregated scores were compared to reference data and presented in a report that ranked the physician in comparison to physicians with a similar practice on different aspects of performance (e.g. similar to CanMEDS dimensions of medical expert, communication, professionalism).(28) The terms 'multisource feedback (MSF)' and 'physician achievement review (PAR)' can essentially be used interchangeably, but for simplicity we use MSF exclusively in this paper.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

We report on the findings of a survey of the entire physician membership in Alberta in 2015. The objectives of this survey were: 1) to discover physicians' actual opinions regarding MSF; 2) to see if different types of physicians (e.g. males versus females, older versus younger) varied in their opinion of MSF.

Methods

This survey, as a component of a larger evaluation and re-design of the College of Physicians and Surgeons of Alberta (CPSA)'s continuing competence programs, received ethical approval from the University of Alberta joint Research Ethics Board. Participants gave informed consent at the time of the survey.

This was an electronic on-line survey sent to all physicians fully licensed to practice medicine in Alberta, Canada in 2015 (n=approximately 9,000). The research team initially developed a series of questions based on the existing literature and from questionnaires and surveys commissioned by the CPSA previously. These questions were then pretested on most of the physicians who worked at the CPSA (a general surgeon, an occupational medicine physician and 5 family physicians). The draft survey was then tested and reviewed by a special advisory committee (the 'pilot' committee) set up to help with piloting, running and interpreting the survey results. The pilot committee consisted of representatives from various physician organizations (Alberta Medical Association, Canadian Medical Association, Alberta Health Services, University of Alberta, University of Calgary), Alberta family physicians (primary care network), a public representative, CPSA Council members, specialist and generalist physicians, CPSA staff and previous leaders of the PAR program. Ultimately the penultimate version of the survey was sent out to three family physician offices (approximately 30 family physicians) to additionally test the process for logging on and completing the online survey. Finally, we sent out an initial request and link via email to the final survey to all physicians in the province (wave 1) followed by a reminder request and link 2 months later (wave 2). We deliberately collected two separate 'waves' or responders to determine whether the demographics of the potential slow or non-responders differed significantly from each other or from the whole population of Alberta physicians. Uniquely in surveys

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

of populations we already knew the exact expected demographic makeup (because the CPSA holds data on all physicians practicing in Alberta) which can help to identify any potential skewing of the responses.

The survey consisted of twelve questions asking the physician how successful MSF was at assessing different areas of physician competence (medical knowledge, clinical skills, communication skills, professionalism, administration, practice management, team functioning,) and how inspiring, motivating and reflective MSF was in their opinion. There was also a text response field for physicians to provide written, open-ended feedback and comments at the end of the questionnaire. Finally, because the survey was anonymous, we asked respondents to provide information on their gender, year of graduation, whether they were a Family Medicine Physician/General Practitioner (FM/GP), if they practiced solo, and whether they obtained their medical degree in Canada (Canadian MD). We used a 10-point modified Likert response scale to answer with three anchors (1= "not at all", 5 "fairly" and 10= "extremely").

Quantitative responses were analyzed using Statistical Package for the Social Sciences (SPSS) version 23 for Mac. Descriptive, bivariate, and multivariate general linear model analysis was performed. The multivariate general linear model was created using the twelve survey questions as the dependent variables, and years in practice, Canadian MD, solo practice, family physician, and gender as the independent variables. Missing data was excluded from the analysis. Any potential interaction terms were tested for and included in the final model. The internal reliability of the survey was measured using Cronbach's alpha.

Qualitative (open-ended) text responses were examined using a thematic inductive qualitative content analysis by a member of the research team with experience conducting qualitative and mixed-methods research and analyses.(29) NVivo™ 11 Pro for Windows qualitative data analysis software was used to assist in grouping the open-ended text data into themes for analysis.

Patient and public involvement

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

A public/patient representative was a member of the survey pilot committee that helped to design the questions and the survey pilot initially. The same committee also reviewed the results and helped to interpret the findings. The results of the survey have already been disseminated to all physicians in Alberta via the monthly newsletter sent out via email to every physician in the province.

Results

2215 physicians responded out of 9021 fully registered Alberta physicians (response rate of 24.6%). Cronbach's alpha for the survey questions was 0.964 indicating very high internal consistency for the survey items.

Wave 1 consisted of 1387 physicians and wave 2 consisted of 828 physicians [Table 1]. The waves were significantly different from each other in terms of age and proportion of Canadian MDs with wave 1 being slightly longer in practice by a mean of four years, and having approximately 6% more Canadian MDs. Compared to the entire physician population, respondents (wave 1 and 2) were slightly longer in practice, more female, more FM, more Canadian MDs, and substantially less solo-practicing physicians [Table 1].

The mean rating for how successful MSF was at assessing a variety of dimensions, varied from a low of 5.03/10 for medical knowledge, to a high of 6.38/10 for professionalism and communication [Table 2]. The overall mean and 95% confidence interval for the twelve MSF survey responses was 5.7 (5.58-5.82) indicating the 'true' mean response is likely to be within about 2% of our value 95 times out of 100.

The multivariate GLM showed that only Canadian MDs had significantly different opinions compared to non-Canadian-trained MDs [Table 3]. All of the other variables were not significant in the model and there were no significant interaction terms.

Qualitative thematic content analysis of the 303 responses to the open-ended text box asking for "Additional Comments and Ideas" regarding MSF, revealed that respondents overwhelmingly regarded MSF as negative. Overarching themes included the opinions that MSF is:

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For Peer Review Only

1) A waste of time and/or resources:

> *To me the [MSF] is an incredible waste of time and money. I do not know one MD who actually pays it much attention, either filling it out or implementing changes.*

2) Irrelevant or useless:

> *I believe that [MSF] is really useless as it currently operates. Major changes required.*

3) Subjective, biased:

> *The responses are biased to some or a large degree as the patients and peers that assess are chosen by the person being reviewed. If the person being reviewed is in an influential position, I'm not sure how accurate the review really is. I try to be as honest as I can when reviewing others, but wonder if that is always the case.*

4) Distressing, stressful:

> *[I] find it [MSF] an extremely stressful exercise that I do every 5 years to maintain my license.*

Discussion

The CPSA research team was genuinely surprised to find such a low/mediocre rating for MSF (on average, 5-6/10), given the previous much more optimistic data from published research studies and non-independent small surveys.(1, 3, 8-10, 23) Results from this research indicate that substantial efforts are required to investigate and potentially improve the utility of MSF in Alberta physicians. Successful MSF programs in business tend to involve considerable investment in training reviewers beforehand, educating participants about the potential benefits of MSF, and providing feedback directly to the participant to help with understanding and facilitate actionable change.(30) These strategies were not used in Alberta with the PAR program, and perhaps those who are considering applying MSF-type programs to physicians, should incorporate the valuable successes and variable lessons learned from the business milieu.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

Furthermore, the substantially lower ratings by Canadian-trained MDs on almost every dimension of MSF were an even greater surprise to the research team. Again, considerable thought and investigation is required to assist in explaining why there is such a marked difference in opinion between these two groups. One could surmise that Canadian-trained MDs are more familiar with performance assessment (and MSF in particular), and can be much more critical in their responses. Conversely, non-Canadian-trained MDs might feel more grateful or appreciative for the opportunity of receiving feedback, particularly if they had not been exposed to similar forms of feedback in their country of medical training. Additionally, non-Canadian-trained MDs might feel more vulnerable, or less comfortable criticizing the CPSA even though the study was anonymous and without consequence.

Survey participation is notoriously difficult in physician groups who are often bombarded with multiple different survey requests on an almost daily basis.(31, 32) Previous surveys from CPSA have met with response rates of about 10% and this may not be unusual even for physician professional organizations (CPSA. CPSA Survey 2014. Unpublished. 2014 May pp. 1–228). Our response rate of 25% was more than double the response rate of any previous survey by the CPSA, indicating that this is clearly an important topic for Alberta physicians. There is growing evidence that high-quality, rich and reliable information can be obtained from surveys with lower response rates (20-25%) and in fact they may be superior to higher response rate surveys.(33-36) One potential reason being that initial 'non-responders' who are 'forced' to answer do not do so in a benevolent way which can seriously undermine the validity of the findings.(37) Secondly there seems to be little if any evidence to support that the typical minimum response rate should be 60-65%.(38, 39) Holbrook et all analyzed results from 81 national surveys (with response rates from 5-54%) and found the extra expense and energy invested in trying to boost response rates may not result in worthwhile improvements in reliability or validity.(40)

Statistically, there was a difference between the different waves of respondents and between the entire population of Alberta physicians in terms of demographics, but given the large sample sizes, it may be very easy to find statistical differences even when the actual difference is small. Given that we know the composition of the responders and all Alberta physicians, we can predict the potential effects of any

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

differences. Looking at the mean differences, wave 1 and 2 participants were very similar in composition between the entire population and the two waves, with one notable exception: there were appreciably less physicians in solo practice who answered the survey (20% versus 30%). Given that the responses from solo physicians were identical to 'non-solo' physicians (with the exception that they rated 'group feedback' measures less importantly for obvious reasons), the validity of the overall results is established. Compared to the entire population of Alberta physicians, respondents were more likely to be female, more likely to be family physicians and more likely to have practiced longer, all of which would not skew the survey responses. There were slightly more Canadian-trained MDs, which might skew all opinions on MSF to be slightly worse than if the entire physician population had responded.

Given the way respondents rated MSF in Alberta, it is difficult to see how MSF could inspire the majority of physicians to make any significant changes to their practice, which is surmised to encourage quality improvement of their performance. The mean score for "inspires reflection in practice" was only fair at 5.69/10. About 30% of physicians did rate the MSF 70% or higher, causing us to question how these physicians might differ from those who rated MSF lower. The statistical analysis found that these higher-MSF-rating physicians are more likely to be non-Canadian-trained MDs. However, the business literature would suggest that these "high rating" respondents may have unique personality variables that are receptive to MSF, such as high levels of emotional stability, extroversion and conscientiousness.(41) Unfortunately, because the survey was anonymous, the personalities of the "high rating" respondents cannot be further explored. Our findings complement research already completed in Alberta, where 72% of a cohort of surgeons taking part in the PAR program indicated initially that they were "contemplating or [had] initiated change on the basis of multisource feedback…".(10) However, after following-up with the cohort three months later, the same authors concluded that "surgeons made few changes in practice in response to feedback data;" postulating that perhaps the surgeons did not value the information they received from MSF because it was not based on surgical outcomes, "the cornerstone of surgical practice."(42)

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Physicians in Alberta have much lower opinions regarding the ability of MSF to measure the dimensions of performance, compared to previous anecdotal reports and published literature to date on this subject. Canadian-trained physicians have a particularly low opinion of MSF for reasons that remain unclear. Further investigation into the conclusions from this research will allow for a richer understanding of these opinions, and offer an opportunity for exploration into the appropriateness of the application of MSF for medical doctors. Individual interviews and/or focus groups with physicians to further explore the themes identified in this inaugural study may lead to an increased understanding of participants' differing opinions of, and experiences with, multisource feedback as it pertains to performance. The results of this survey offer a serious challenge to the effectiveness of a program that is intended to promote self-reflection and performance improvement in physicians.

For peer review only

Table 1: Demographics of responders and non-responders

	Wave 1 (n=1387)	Wave 2 (n=828)	Wave1 versus Wave2 significance	All (n=9021)	Wave 1+2 versus All significance
Years in practice (sd)	26.0 (11.7)	22.9 (11.8)	<0.001	22.0 (12.2)	<0.001
Gender (% female)	38.8	41.4	NS (0.37)	37.2	0.02
Family physician (%)	51.8	57.1	NS (0.07)	48.5	0.001
Solo practice (%)	19.8	18.8	NS (0.67)	30.5	<0.001
Canadian MD (%)	73.8	67.3	0.02	66.6	<0.001

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2: Mean rating for question "Please rate how successful the existing [multisource feedback] program is in assessing the following dimensions. (1-10 where 1=not at all, 5=fairly, and 10= extremely)"

Dimension	Mean Score	Standard error
Medical knowledge	5.03	0.06
Clinical skills	5.17	0.06
Communication skills	6.38	0.06
Practice administration	5.36	0.06
Patient management	5.68	0.06
Adherence to standards of practice	5.34	0.06
Professionalism	6.38	0.06
Team functioning	6.15	0.06
Easy to participate	6.20	0.07
Inspires reflection in practice	5.69	0.07
Motivates clinical practice improve	5.58	0.07
Provide a learning opportunity	5.47	0.07
Overall	5.70	0.06

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3: Mean rating for question "Please rate how successful the existing [multisource feedback] program is in assessing the following dimensions. (1-10 where 1=not at all, 5=fairly, and 10= extremely)" stratified by Canadian or Non-Canadian trained MD.

Dimension	Non-Canadian trained MD	Canadian trained MD	Multivariate analysis (p value)
Medical knowledge	5.71	4.74	<0.0001
Clinical skills	5.72	4.92	<0.0001
Communication skills	6.61	6.31	NS
Practice administration	6.04	5.08	<0.0001
Patient management	6.34	5.41	<0.0001
Adherence to standards of practice	6.11	4.98	<0.0001
Professionalism	6.73	6.30	NS
Team functioning	6.52	6.02	NS
Easy to participate	6.79	6.03	<0.0001
Inspires reflection in practice	6.35	5.43	<0.0001
Motivates clinical practice improve	6.29	5.31	<0.0001
Provide a learning opportunity	6.1	5.26	<0.0001
Overall	6.19	5.48	

NS = Not significant at 0.05 level

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Acknowledgements

Tanya Northfield for doing all the hard work and her endless patience with all of us. Dr. Ernie Schuster for the initial idea and help early on with the project.

Conflicts of interest

All the authors work for CPSA who ran the physician achievement review (PAR) program, a MSF based assessment program for 18 years in Alberta.

Author contributions:

NA and KM had the initial concept. NA, NC and EJ designed the study. NA and NC collected the data, performed the quantitative and qualitative analyses. NA, NC, EJ and KM reviewed and interpreted the results. NA drafted the manuscript. All the authors wrote and edited and revised the final versions of the manuscript.

Data availability statement: All data requests should be submitted to the corresponding author for consideration. Access to anonymized data will not be possible

Funding: None

For peer review only

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in BMJ Open and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

References

1. Hall W, Violato C, Lewkonia R, Lockyer J, Fidler H, Toews J, et al. Assessment of physician performance in Alberta: the physician achievement review. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 1999;161(1):52-7.

2. Bracken DW, Rose DS, Church AH. The evolution and devolution of 360 feedback. Industrial and Organizational Psychology. 2016;9(4):761-94.

3. Violato C, Marini A, Toews J, Lockyer J, Fidler H. Feasibility and psychometric properties of using peers, consulting physicians, co-workers, and patients to assess physicians. Academic medicine : journal of the Association of American Medical Colleges. 1997;72(10 Suppl 1):S82-4.

4. Campbell JL, Roberts M, Wright C, Hill J, Greco M, Taylor M, et al. Factors associated with variability in the assessment of UK doctors' professionalism: analysis of survey results. BMJ (Clinical research ed). 2011;343(oct27 2):d6212-d.

5. Mount M, Judge T, Scullen S. Trait, rater and level effects in 360-degree performance ratings. Personnel Psychology. 1998;51(3):557-76.

6. Ansari A, Donnon T, Al Khalifa K, Darwish A, Violato C. The construct and criterion validity of the multi-source feedback process to assess physician performance: a meta-analysis. Advances in Medical Education and Practice. 2014;5:39.

7. Donnon T, Al Ansari A, Al Alawi S, Violato C. The reliability, validity, and feasibility of multisource feedback physician assessment: a systematic review. Academic medicine : journal of the Association of American Medical Colleges. 2014;89(3):511-6.

8. Lockyer J. Multisource feedback in the assessment of physician competencies. The Journal of continuing education in the health professions. 2003;23(1):4-12.

9. Lockyer JM, Violato C, Fidler H. A multi source feedback program for anesthesiologists. Canadian journal of anaesthesia = Journal canadien d'anesthésie. 2006;53(1):33-9.

10. Violato C, Lockyer J, Fidler H. Multisource feedback: a method of assessing surgical practice. BMJ (Clinical research ed). 2003;326(7388):546-8.

For peer review only

11. Burford B, Illing J, Kergon C, Morrow G, Livingston M. User perceptions of multi-source feedback tools for junior doctors. Medical education. 2010;44(2):165-76.

12. Overeem K, Wollersheim H, Driessen E, Lombarts K, van de Ven G, Grol R, et al. Doctors' perceptions of why 360-degree feedback does (not) work: a qualitative study. Medical education. 2009;43(9):874-82.

13. Yama BA, Hodgins M, Boydell K, Schwartz SB. A qualitative exploration: questioning multisource feedback in residency education. BMC Medical Education. 2018;18(1):170.

14. Ferguson J, Wakeling J, Bowie P. Factors influencing the effectiveness of multisource feedback in improving the professional practice of medical doctors: a systematic review. BMC Medical Education. 2014;14(1):1-12.

15. Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. BMJ (Clinical research ed). 2010;341(sep24 1):c5064-c.

16. Brinkman WB, Geraghty SR, Lanphear BP, Khoury JC, Gonzalez del Rey JA, Dewitt TG, et al. Effect of multisource feedback on resident communication skills and professionalism: a randomized controlled trial. Archives of pediatrics & adolescent medicine. 2007;161(1):44-9.

17. Ilgen D, Davis C. Bearing Bad News: Reactions to Negative Performance Feedback. Applied Psychology. 2000;49(3):550-65.

18. Nowack KM, Mashihi S. Evidence-based answers to 15 questions about leveraging 360-degree feedback. Consulting Psychology Journal: Practice and Research. 2012;64(3):157-82.

19. Kluger AN, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin. 1996;119(2):254-84.

20. Peiperl MA. Getting 360 degrees feedback right. Harvard business review. 2001;79(1):142-7- 77.

21. Doran GT. There's a S.M.A.R.T. way to write management's goals and objectives. Manage Rev. 1981;70(11):35-6.

22. Atwater LE, Brett JF, Charles AC. Multisource feedback: Lessons learned and implications for practice. Human Resource Management. 2007;46(2):285-307.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

23. Sargeant JM, Mann KV, Ferrier SN, Langille DB, Muirhead PD, Hayes VM, et al. Responses of rural family physicians and their colleague and coworker raters to a multi-source feedback process: a pilot study. Academic medicine : journal of the Association of American Medical Colleges. 2003;78(10 Suppl):S42-4.

24. Ramsey PG, Carline JD, Blank LL, Wenrich MD. Feasibility of hospital-based use of peer ratings to evaluate the performances of practicing physicians. Academic medicine : journal of the Association of American Medical Colleges. 1996;71(4):364-70.

25. Goodyear HM, Lakshminarayana I, Wall D, Bindal T. A multisource feedback tool to assess ward round leadership skills of senior paediatric trainees: (2) Testing reliability and practicability. Postgraduate medical journal. 2015;91(1075):268-73.

26. Health Professions Act, (Jul 09, 2018).

27. Lockyer JM, Violato C, Fidler HM. What multisource feedback factors influence physician self-assessments? A five-year longitudinal study. Academic medicine : journal of the Association of American Medical Colleges. 2007;82(10 Suppl):S77-80.

28. Frank J, Snell L, Sherbino J. CanMEDS 2015 Physician Competency Framework. Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.

29. Elo S, Kyngäs H. The qualitative content analysis process. Journal of advanced nursing. 2008;62(1):107-15.

30. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. Medical teacher. 2006;28(7):e185-91.

31. Cunningham CT, Quan H, Hemmelgarn B, Noseworthy T, Beck CA, Dixon E, et al. Exploring physician specialist response rates to web-based surveys. BMC medical research methodology. 2015;15(1):32.

32. Tambor ES, Chase GA, Faden RR, Geller G, Hofman KJ, Holtzman NA. Improving response rates through incentive and follow-up: the effect on a survey of physicians' knowledge of genetics. American journal of public health. 2011;83(11):1599-603.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

33. Choung RS, Locke GR, Schleck CD, Ziegenfuss JY, Beebe TJ, Zinsmeister AR, et al. A low response rate does not necessarily indicate non-response bias in gastroenterology survey research: a population-based study. Journal of Public Health. 2012;21(1):87-95.

34. Curtin R, Presser S, Singer E. The effects of response rate changes on the index of consumer sentiment. Public Opinion Quarterly. 2000;64(4):413-28.

35. Keeter S, Kennedy C, Dimock M. Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. Public Opinion Quarterly. 2006;70:759-79.

36. Visser PS, Krosnick JA, Marquette J. Mail surveys for election forecasting? An evaluation of the columbus dispatch poll. Public Opinion Quarterly. 1996;60:181-227.

37. Decieux JP, A M, KM N, Sischka P. Implementation of the forced answering option within online surveys: Do higher item response rates come at the expense of participation and answer quality? Psihologija. 2015;48(4):311-26.

38. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Public Opinion Quarterly. 2006;5(70):646-75.

39. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? Epidemiology (Cambridge, Mass). 2006;17(4):413-8.

40. Holbrook A, Krosnick J, Pfent A. The causes and consequences of response rates in surveys by the news media and government contractor survey research firms. In: Lepkowski JM, editor. Advances in telephone survey methodology. New York: Wiley; 2007. p. 1-92.

41. Smither JW, London M, Reilly RR. Does Performance Improve Following Multisource Feedback? a Theoretical Model, Meta-Analysis, and Review of Empirical Findings. Personnel Psychology. 2005;58(1):33-66.

42. Lockyer J, Violato C, Fidler H. Likelihood of change: a study assessing surgeon use of multisource feedback data. Teaching and learning in medicine. 2003;15(3):168-74.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
	page 1,2	(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	p 5,6 2	Explain the scientific background and rationale for the investigation being reported
Objectives	p 7 3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	p 7 4	Present key elements of study design early in the paper
Setting	p 7 5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	p 7 6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants
Variables	p 8 7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	p 7,8 8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	p 7,8 9	Describe any efforts to address potential sources of bias
Study size	p 7 10	Explain how the study size was arrived at
Quantitative variables	p 8 11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	p 8 12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
	p 9 Table 1	(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	p 9 14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
	Table 1	(b) Indicate number of participants with missing data for each variable of interest
Outcome data	p 9,10 15*	Report numbers of outcome events or summary measures
Main results	p 9,10 16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
	Table 2,3	(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open: first published as 10.1136/bmjopen-2020-037610 on 19 July 2020. Downloaded from http://bmjopen.bmj.com/ on April 9, 2024 by guest. Protected by copyright.

Discussion			
Key results	p10	18	Summarise key results with reference to study objectives
Limitations	p11	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	p10-12	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	p13	21	Discuss the generalisability (external validity) of the study results
Other information			
Funding	p17	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open: first published as 10.1136/bmjopen-2020-037610 on 19 July 2020. Downloaded from http://bmjopen.bmj.com/ on April 9, 2024 by guest. Protected by copyright.

BMJ Open

Survey of physician attitudes to using multisource feedback for competence assessment in Alberta

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-037610.R2
Article Type:	Original research
Date Submitted by the Author:	01-Jun-2020
Complete List of Authors:	Ashworth, Nigel; University of Alberta Faculty of Medicine and Dentistry, Medicine; College of Physicians & Surgeons of Alberta, Continuing Competence Kain, Nicole; College of Physicians & Surgeons of Alberta, Continuing Competence; University of Alberta Faculty of Medicine and Dentistry, Medicine Jess, Ed; The College of Physicians & Surgeons of Alberta, Continuing Competence Mazurek, Karen; The College of Physicians & Surgeons of Alberta, Continuing Competence
Primary Subject Heading:	Medical education and training
Secondary Subject Heading:	Qualitative research, Communication
Keywords:	EDUCATION & TRAINING (see Medical Education & Training), Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, MEDICAL EDUCATION & TRAINING

SCHOLARONE™
Manuscripts

BMJ Open: first published as 10.1136/bmjopen-2020-037610 on 19 July 2020. Downloaded from http://bmjopen.bmj.com/ on April 9, 2024 by guest. Protected by copyright.

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

For peer review only

Survey of physician attitudes to using multisource feedback for competence assessment in Alberta

Nigel Ashworth, MB ChB., MSc.[1,2], Nicole Kain, PhD.[1,2], Ed Jess, BA.[2] and Karen Mazurek, MD.[2]

[1] Department of Medicine, University of Alberta, Edmonton; [2] College of Physicians and Surgeons of Alberta, Edmonton

No funding, conflicts or competing interests to declare from any author.

Corresponding author: ashworth@ualberta.ca

Word Count: 3034

Keywords: Physicians, MSF, Multisource, Feedback, Attitudes

For peer review only

Abstract

Background: The use of multisource feedback (MSF) for assessing physician performance is widespread and rapidly growing. Findings from early very small research studies using highly selected participants suggest high levels of satisfaction and support. However, after nearly two decades of experience using MSF to evaluate all physicians in Alberta we are skeptical of this.

Objectives: To determine physicians' actual opinions of MSF using the entire physician population of Alberta, Canada

Design: On-line survey

Setting: Alberta, Canada

Participants: All physicians with a full license to practice in Alberta in 2015.

Interventions: All participants were asked to grade how well they thought MSF was at assessing various aspects of physician performance using a 10-point Likert-type scale. There was also a text response field for written comments.

Outcomes: Mean responses to quantitative questions. Qualitative content and thematic analysis of open-ended text responses.

We analyzed the data using SPSS23 and NVivo11 and built a multivariate model highlighting the predictors of high and low opinions of MSF.

Results: Survey response rate was high for physicians with 2215 responses (25%). The mean rating for how successful MSF was at assessing a variety of dimensions, varied from a low of 5.03/10 for medical knowledge, to a high of 6.38/10 for professionalism and communication. Canadian trained MDs rated MSF significantly lower on every dimension by approximately 20% compared to non-Canadian trained MDs.

Conclusions: Alberta physicians have much lower opinions about the ability of MSF to measure any dimension of their performance than what has been suggested in the literature. Canadian-trained MDs

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

have a particularly low opinion of MSF for reasons that remain unclear. The results of this survey offer a serious challenge to the effectiveness of a program that is designed to promote self-reflection and performance improvement.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Strengths and limitations of this study

- MSF is of considerable interest to physicians Worldwide given it is increasingly being used as part of QI and QA assessments
- This is the largest study of its kind to date
- We employed a mixed methods analysis that allowed for open-ended text responses to give in-depth insight into the quantitative responses
- We used multivariate statistical methods to analyze the quantitative data
- Survey was anonymous but some physicians may still have been uncomfortable giving honest answers to the medical regulator

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Introduction

Multisource feedback (MSF) also known as 360-degree or multi-source or multi-rater feedback has been used as a tool for quality improvement and quality assurance in healthcare for over 20 years and in the business world for even longer.(1, 2) One advantage of MSF is that it can provide a physician with a much broader range of personal feedback (e.g. medical expert, communication, administration, professionalism) from many different sources (patient, co-worker, peer, self, leader) than might be obtained traditionally.(1, 3) In addition, because feedback from many patients and multiple peers and co-workers is typically obtained and aggregated, then this should average out any more extreme outlier views that assessors might express, hopefully giving a more balanced overall view.(4, 5) The overall psychometric properties, in terms of the reliability, validity and feasibility of the tool have been reported to be good.(3, 6, 7)

However, despite the widespread and rapidly growing use of MSF for both formative and summative assessments of physician competence and performance, the actual views of physicians regarding MSF have not been well explored. Previous studies of physician attitudes to MSF in Canada and Worldwide have shown very high rates of satisfaction with MSF, for example in one paper from Alberta 100% of 24 physicians felt that "[MSF] was a helpful educational exercise…".(3) Another study of 308 physician volunteers found 2/3 of those who responded to a post study satisfaction questionnaire (numbers not given) said they were contemplating or had made changes to their practice.(1) Other studies have typically surveyed similar small numbers of highly self-selected physicians.(8-10) A survey of 249 junior doctor attitudes in the UK found positive attitudes to MSF but low perceived opinion about the effectiveness.(11) This is surprising, given that a key assumption of MSF is that the 'weak' areas or opportunities for practice improvement identified will encourage physicians to make changes and improve in those areas. If the physician being assessed is not supportive of the process, then one can hardly claim that the individual will be likely to make changes. A qualitative study of consultant physicians in the Netherlands identified 'lack of openness and constructive feedback' as major barriers for the success of a MSF program.(12) Similarly, a qualitative study of opinions about MSF amongst

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

allied health staff and pediatric residents in Canada showed strong interest in the concept of MSF but significant potential barriers to success, such as poorly defined roles and responsibilities, perceptions of expertise, hostile hospital culture and negative interprofessionalism and power dynamics.(13). Only 1/16 studies looking at the effectiveness of MSF in physicians identified any significant positive change in actual behavior and in that study the treatment group also received "…a tailored coaching session to assist in identifying their strengths and weaknesses and in setting specific behavioural goals".(14-16) Furthermore, it is well-established in the business literature that MSF programs can cause significant negative effects with a decreased performance in up to one-third of all MSF programs due in part to an "exacerbation [of] bureaucracy, heightening [of] political tensions, and consumption [of] enormous numbers of hours".(17-20) Also giving good feedback is not easy, ideally the source should be credible and the information should be 'SMART': specific, measurable, attainable, relevant, and time-bound.(21) Additionally, significant anxiety and distress can be imposed upon individual physicians who are told they are in the 'bottom 10%.(22) Several small research projects of selected volunteer physicians who were used as participants to test new MSF programs or components of programs, found high rates of satisfaction of around 70%.(3, 9, 23-25) However, these results can hardly be said to represent the views of the majority of physicians. Such contradictory findings in the literature have created an urgent need to better understand physicians' attitudes and opinions regarding MSF.

Since 1999, all physicians in Alberta were mandated to undergo a competence/performance assessment at least once every five years, in accordance with Alberta's Health Professions Act.(26) The assessment, also known as the Physician Achievement Review (PAR) program, comprised a MSF assessment involving questionnaires from eight peer physicians (colleagues), eight co-workers, (e.g. nurses, receptionists, physiotherapists), twenty-five patients and a self-evaluation.(27) The physician's aggregated scores were compared to reference data and presented in a report that ranked the physician in comparison to physicians with a similar practice on different aspects of performance (e.g. similar to CanMEDS dimensions of medical expert, communication, professionalism).(28) The terms 'multisource feedback (MSF)' and 'physician achievement review (PAR)' can essentially be used interchangeably, but for simplicity we use MSF exclusively in this paper.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

We report on the findings of a survey of the entire physician membership in Alberta in 2015. The objectives of this survey were: 1) to discover physicians' actual opinions regarding MSF; 2) to see if different types of physicians (e.g. males versus females, older versus younger) varied in their opinion of MSF.

Methods

This survey, as a component of a larger evaluation and re-design of the College of Physicians and Surgeons of Alberta (CPSA)'s continuing competence programs, received ethical approval from the University of Alberta joint Research Ethics Board. Participants gave informed consent at the time of the survey.

This was an electronic on-line survey sent via email to all physicians fully licensed to practice medicine in Alberta, Canada in 2015 (n=approximately 9,000). The research team initially developed a series of questions based on the existing literature and from questionnaires and surveys commissioned by the CPSA previously. These questions were then pretested on most of the physicians who worked at the CPSA (a general surgeon, an occupational medicine physician and 5 family medicine physicians). The draft survey was then tested and reviewed by a special advisory committee (the 'pilot' committee) set up to help with piloting, running and interpreting the survey results. The pilot committee consisted of representatives from various physician organizations (Alberta Medical Association, Canadian Medical Association, Alberta Health Services, University of Alberta, University of Calgary), Alberta family physicians (primary care network), a public representative, CPSA Council members, specialist and generalist physicians, CPSA staff and previous leaders of the PAR program. The penultimate version of the survey was sent out to three family physician offices (approximately 30 family physicians) to additionally test the process for logging on and completing the online survey. Finally, we sent out an initial request and link via email to the final survey to all physicians in the province (wave 1) followed by a reminder request and link 2 months later (wave 2). We deliberately collected two separate 'waves' or responders to determine whether the demographics of the potential slow or non-responders differed significantly from each other or from the whole population of Alberta physicians. Uniquely in surveys

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

of populations we already knew the exact expected demographic makeup (because the CPSA holds data on all physicians practicing in Alberta) which can help to identify any potential skewing of the responses.

The survey consisted of twelve questions asking the physician how successful MSF was at assessing different areas of physician competence (medical knowledge, clinical skills, communication skills, professionalism, administration, practice management, team functioning,) and how inspiring, motivating and reflective MSF was in their opinion. There was also a text response field for physicians to provide written, open-ended feedback and comments at the end of the questionnaire. Finally, because the survey was anonymous, we asked respondents to provide information on their gender, year of graduation, whether they were a Family Medicine Physician/General Practitioner (FM/GP), if they practiced solo (compared to practicing in a group), and whether they obtained their medical degree in Canada (Canadian MD). We used a 10-point modified Likert response scale to answer with three anchors (1= "not at all", 5 "fairly" and 10= "extremely").

Quantitative responses were analyzed using Statistical Package for the Social Sciences (SPSS) version 23 for Mac. Descriptive, bivariate, and multivariate general linear model analysis was performed. The multivariate general linear model was created using the twelve survey questions as the dependent variables, and years in practice, Canadian MD, solo practice, family physician, and gender as the independent variables. Missing data was excluded from the analysis. Any potential interaction terms were tested for and included in the final model. The internal reliability of the survey was measured using Cronbach's alpha.

Qualitative (open-ended) text responses were examined using a thematic inductive qualitative content analysis by a member of the research team with experience conducting qualitative and mixed-methods research and analyses.(29) NVivo™ 11 Pro for Windows qualitative data analysis software was used to assist in grouping the open-ended text data into themes for analysis.

Patient and public involvement

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

A public/patient representative was a member of the survey pilot committee that helped to design the questions and the survey pilot initially. The same committee also reviewed the results and helped to interpret the findings. The results of the survey have already been disseminated to all physicians in Alberta via the monthly newsletter sent out via email to every physician in the province.

Results

2215 physicians responded out of 9021 fully registered Alberta physicians (response rate of 24.6%). Cronbach's alpha for the survey questions was 0.964 indicating very high internal consistency for the survey items.

Wave 1 consisted of 1387 physicians and wave 2 consisted of 828 physicians [Table 1]. The waves were significantly different from each other in terms of age and proportion of Canadian MDs with wave 1 being slightly longer in practice by a mean of four years, and having approximately 6% more Canadian MDs. Compared to the entire physician population, respondents (wave 1 and 2) were slightly longer in practice, more female, more FM, more Canadian MDs, and substantially less solo-practicing physicians [Table 1].

The mean rating for how successful MSF was at assessing a variety of dimensions, varied from a low of 5.03/10 for medical knowledge, to a high of 6.38/10 for professionalism and communication [Table 2]. The overall mean and 95% confidence interval for the twelve MSF survey responses was 5.7 (5.58-5.82) indicating the 'true' mean response is likely to be within about 2% of our value 95 times out of 100.

The multivariate GLM showed that only Canadian MDs had significantly different opinions compared to non-Canadian-trained MDs [Table 3]. All of the other variables were not significant in the model and there were no significant interaction terms.

Qualitative thematic content analysis of the 303 responses to the open-ended text box asking for "Additional Comments and Ideas" regarding MSF, revealed that respondents overwhelmingly regarded MSF as negative. Overarching themes included the opinions that MSF is:

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For Peer Review Only

1) A waste of time and/or resources:

> *To me the [MSF] is an incredible waste of time and money. I do not know one MD who actually pays it much attention, either filling it out or implementing changes.*

2) Irrelevant or useless:

> *I believe that [MSF] is really useless as it currently operates. Major changes required.*

3) Subjective, biased:

> *The responses are biased to some or a large degree as the patients and peers that assess are chosen by the person being reviewed. If the person being reviewed is in an influential position, I'm not sure how accurate the review really is. I try to be as honest as I can when reviewing others, but wonder if that is always the case.*

4) Distressing, stressful:

> *[I] find it [MSF] an extremely stressful exercise that I do every 5 years to maintain my license.*

Discussion

The CPSA research team was genuinely surprised to find such a low/mediocre rating for MSF (on average, 5-6/10), given the previous much more optimistic data from published research studies and non-independent small surveys.(1, 3, 8-10, 23) Results from this research indicate that substantial efforts are required to investigate and potentially improve the utility of MSF in Alberta physicians. Successful MSF programs in business tend to involve considerable investment in training reviewers beforehand, educating participants about the potential benefits of MSF, and providing feedback directly to the participant to help with understanding and facilitate actionable change.(30) These strategies were not used in Alberta with the PAR program, and perhaps those who are considering applying MSF-type programs to physicians, should incorporate the valuable successes and variable lessons learned from the business milieu.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

Furthermore, the substantially lower ratings by Canadian-trained MDs on almost every dimension of MSF were an even greater surprise to the research team. Again, considerable thought and investigation is required to assist in explaining why there is such a marked difference in opinion between these two groups. One could surmise that Canadian-trained MDs are more familiar with performance assessment (and MSF in particular), and can be much more critical in their responses. Conversely, non-Canadian-trained MDs might feel more grateful or appreciative for the opportunity of receiving feedback, particularly if they had not been exposed to similar forms of feedback in their country of medical training. Additionally, non-Canadian-trained MDs might feel more vulnerable, or less comfortable criticizing the CPSA even though the study was anonymous and without consequence. Our survey was generic for all types of physicians and surgeons and it may be that a discipline specific questionnaire might be more valid to allow for different working environments that various physicians experience. Also, the qualitative piece of our study was limited in its scope and not detailed enough to provide more than a tantalizing glimpse into the reasons for these results. Future more in-depth qualitative research, such as individual interviews with physicians or focus groups with respondents, is needed to investigate the themes identified in this work further. Some of this work has already commenced, in a separate but related project that explored the experiential knowledge of physician-assessors in identifying potential risk and support factors to physician performance. (31)

Survey participation is notoriously difficult in physician groups who are often bombarded with multiple different survey requests on an almost daily basis.(32,33) Previous surveys from CPSA have met with response rates of about 10% and this may not be unusual even for physician professional organizations (CPSA. CPSA Survey 2014. Unpublished. 2014 May pp. 1–228). Our response rate of 25% was more than double the response rate of any previous survey by the CPSA, indicating that this is clearly an important topic for Alberta physicians. There is growing evidence that high-quality, rich and reliable information can be obtained from surveys with lower response rates (20-25%) and in fact they may be superior to higher response rate surveys.(34-37) One potential reason being that initial 'non-responders' who are 'forced' to answer do not do so in a benevolent way which can seriously undermine the validity of the findings.(38) Secondly there seems to be little if any evidence to support that the typical

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

minimum response rate should be 60-65%.(39,40) Holbrook et all analyzed results from 81 national surveys (with response rates from 5-54%) and found the extra expense and energy invested in trying to boost response rates may not result in worthwhile improvements in reliability or validity.(41)

Statistically, there was a difference between the different waves of respondents and between the entire population of Alberta physicians in terms of demographics, but given the large sample sizes, it may be very easy to find statistical differences even when the actual difference is small. Given that we know the composition of the responders and all Alberta physicians, we can predict the potential effects of any differences. Looking at the mean differences, wave 1 and 2 participants were very similar in composition between the entire population and the two waves, with one notable exception: there were appreciably less physicians in solo practice who answered the survey (20% versus 30%). Given that the responses from solo physicians were identical to 'non-solo' physicians (with the exception that they rated 'group feedback' measures less importantly for obvious reasons), the validity of the overall results is established. Compared to the entire population of Alberta physicians, respondents were more likely to be female, more likely to be family physicians and more likely to have practiced longer, all of which would not skew the survey responses. There were slightly more Canadian-trained MDs, which might skew all opinions on MSF to be slightly worse than if the entire physician population had responded.

Given the way respondents rated MSF in Alberta, it is difficult to see how MSF could inspire the majority of physicians to make any significant changes to their practice, which is surmised to encourage quality improvement of their performance. The mean score for "inspires reflection in practice" was only fair at 5.69/10. About 30% of physicians did rate the MSF 70% or higher, causing us to question how these physicians might differ from those who rated MSF lower. The statistical analysis found that these higher-MSF-rating physicians are more likely to be non-Canadian-trained MDs. However, the business literature would suggest that these "high rating" respondents may have unique personality variables that are receptive to MSF, such as high levels of emotional stability, extroversion and conscientiousness.(42) Unfortunately, because the survey was anonymous, the personalities of the "high rating" respondents cannot be further explored. Our findings complement research already

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

For peer review only

completed in Alberta, where 72% of a cohort of surgeons taking part in the PAR program indicated initially that they were "contemplating or [had] initiated change on the basis of multisource feedback…".(10) However, after following-up with the cohort three months later, the same authors concluded that "surgeons made few changes in practice in response to feedback data;" postulating that perhaps the surgeons did not value the information they received from MSF because it was not based on surgical outcomes, "the cornerstone of surgical practice."(43)

Physicians in Alberta have much lower opinions regarding the ability of MSF to measure the dimensions of performance, compared to previous anecdotal reports and published literature to date on this subject. Canadian-trained physicians have a particularly low opinion of MSF for reasons that remain unclear. Further investigation into the conclusions from this research will allow for a richer understanding of these opinions, and offer an opportunity for exploration into the appropriateness of the application of MSF for medical doctors. Individual interviews and/or focus groups with physicians to further explore the themes identified in this inaugural study may lead to an increased understanding of participants' differing opinions of, and experiences with, multisource feedback as it pertains to performance. The results of this survey offer a serious challenge to the effectiveness of a program that is intended to promote self-reflection and performance improvement in physicians.

For peer review only

Table 1: Demographics of responders and non-responders

	Wave 1 (n=1387)	Wave 2 (n=828)	Wave1 versus Wave2 significance	All (n=9021)	Wave 1+2 versus All significance
Years in practice (sd)	26.0 (11.7)	22.9 (11.8)	<0.001	22.0 (12.2)	<0.001
Gender (% female)	38.8	41.4	NS (0.37)	37.2	0.02
Family physician (%)	51.8	57.1	NS (0.07)	48.5	0.001
Solo practice (%)	19.8	18.8	NS (0.67)	30.5	<0.001
Canadian MD (%)	73.8	67.3	0.02	66.6	<0.001

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 2: Mean rating for question "Please rate how successful the existing [multisource feedback] program is in assessing the following dimensions. (1-10 where 1=not at all, 5=fairly, and 10= extremely)"

Dimension	Mean Score	Standard error
Medical knowledge	5.03	0.06
Clinical skills	5.17	0.06
Communication skills	6.38	0.06
Practice administration	5.36	0.06
Patient management	5.68	0.06
Adherence to standards of practice	5.34	0.06
Professionalism	6.38	0.06
Team functioning	6.15	0.06
Easy to participate	6.20	0.07
Inspires reflection in practice	5.69	0.07
Motivates clinical practice improve	5.58	0.07
Provide a learning opportunity	5.47	0.07
Overall	5.70	0.06

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Table 3: Mean rating for question "Please rate how successful the existing [multisource feedback] program is in assessing the following dimensions. (1-10 where 1=not at all, 5=fairly, and 10= extremely)" stratified by Canadian or Non-Canadian trained MD.

Dimension	Non-Canadian trained MD	Canadian trained MD	Multivariate analysis (p value)
Medical knowledge	5.71	4.74	<0.0001
Clinical skills	5.72	4.92	<0.0001
Communication skills	6.61	6.31	NS
Practice administration	6.04	5.08	<0.0001
Patient management	6.34	5.41	<0.0001
Adherence to standards of practice	6.11	4.98	<0.0001
Professionalism	6.73	6.30	NS
Team functioning	6.52	6.02	NS
Easy to participate	6.79	6.03	<0.0001
Inspires reflection in practice	6.35	5.43	<0.0001
Motivates clinical practice improve	6.29	5.31	<0.0001
Provide a learning opportunity	6.1	5.26	<0.0001
Overall	6.19	5.48	

NS = Not significant at 0.05 level

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

Acknowledgements

Tanya Northfield for doing all the hard work and her endless patience with all of us. Dr. Ernie Schuster for the initial idea and help early on with the project.

Conflicts of interest

All the authors work for CPSA who ran the physician achievement review (PAR) program, a MSF based assessment program for 18 years in Alberta.

Author contributions:

NA and KM had the initial concept. NA, NK and EJ designed the study. NA and NK collected the data, performed the quantitative and qualitative analyses. NA, NK, EJ and KM reviewed and interpreted the results. NA drafted the manuscript. All the authors wrote and edited and revised the final versions of the manuscript.

Data availability statement: All data requests should be submitted to the corresponding author for consideration. Access to anonymized data will not be possible

Funding: None

For peer review only

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in BMJ Open and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

References

1. Hall W, Violato C, Lewkonia R, Lockyer J, Fidler H, Toews J, et al. Assessment of physician performance in Alberta: the physician achievement review. CMAJ. 1999;161(1):52-7.

2. Bracken DW, Rose DS, Church AH. The evolution and devolution of 360 feedback. Industrial and Organizational Psychology. 2016;9(4):761-94.

3. Violato C, Marini A, Toews J, Lockyer J, Fidler H. Feasibility and psychometric properties of using peers, consulting physicians, co-workers, and patients to assess physicians. Acad Med. 1997;72(10 Suppl 1):S82-4.

4. Campbell JL, Roberts M, Wright C, Hill J, Greco M, Taylor M, et al. Factors associated with variability in the assessment of UK doctors' professionalism: analysis of survey results. BMJ. 2011;343:d6212.

5. Mount M, Judge T, Scullen S. Trait, rater and level effects in 360-degree performance ratings. Personnel Psychology. 1998;51(3):557-76.

6. Ansari Al, Donnon T, Al Khalifa K, Darwish A, Violato C. The construct and criterion validity of the multi-source feedback process to assess physician performance: a meta-analysis. AMEP. 2014;5:39.

7. Donnon T, Al Ansari A, Al Alawi S, Violato C. The reliability, validity, and feasibility of multisource feedback physician assessment: a systematic review. Acad Med. 2014;89(3):511-6.

8. Lockyer J. Multisource feedback in the assessment of physician competencies. J Contin Educ Health Prof. 2003;23(1):4-12.

9. Lockyer JM, Violato C, Fidler H. A multi source feedback program for anesthesiologists. Can J Anaesth. 2006;53(1):33-9.

10. Violato C, Lockyer J, Fidler H. Multisource feedback: a method of assessing surgical practice. BMJ. 2003;326(7388):546-8.

11. Burford B, Illing J, Kergon C, Morrow G, Livingston M. User perceptions of multi-source feedback tools for junior doctors. Med Educ. 2010;44(2):165-76.

For peer review only

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

12. Overeem K, Wollersheim H, Driessen E, Lombarts K, van de Ven G, Grol R, et al. Doctors' perceptions of why 360-degree feedback does (not) work: a qualitative study. Med Educ. 2009;43(9):874-82.

13. Yama BA, Hodgins M, Boydell K, Schwartz SB. A qualitative exploration: questioning multisource feedback in residency education. BMC Med Educ. 2018;18(1):170.

14. Ferguson J, Wakeling J, Bowie P. Factors influencing the effectiveness of multisource feedback in improving the professional practice of medical doctors: a systematic review. BMC Med Educ. 2014;14(1):1-12.

15. Miller A, Archer J. Impact of workplace based assessment on doctors' education and performance: a systematic review. BMJ. 2010;341:c5064.

16. Brinkman WB, Geraghty SR, Lanphear BP, Khoury JC, Gonzalez del Rey JA, Dewitt TG, et al. Effect of multisource feedback on resident communication skills and professionalism: a randomized controlled trial. Arch Pediatr Adolesc Med. 2007;161(1):44-9.

17. Ilgen D, Davis C. Bearing bad news: reactions to negative performance feedback. Applied Psychology. 2000;49(3):550-65.

18. Nowack KM, Mashihi S. Evidence-based answers to 15 questions about leveraging 360-degree feedback. Consulting Psychology Journal: Practice and Research. 2012;64(3):157-82.

19. Kluger AN, DeNisi A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin. 1996;119(2):254-84.

20. Peiperl MA. Getting 360 degrees feedback right. Harv Bus Rev. 2001;79(1):142-7- 77.

21. Doran GT. There's a S.M.A.R.T. way to write management's goals and objectives. Manage Rev. 1981;70(11):35-6.

22. Atwater LE, Brett JF, Charles AC. Multisource feedback: Lessons learned and implications for practice. Hum Resour Manage. 2007;46(2):285-307.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

23. Sargeant JM, Mann KV, Ferrier SN, Langille DB, Muirhead PD, Hayes VM, et al. Responses of rural family physicians and their colleague and coworker raters to a multi-source feedback process: a pilot study. Acad Med. 2003;78(10 Suppl):S42-4.

24. Ramsey PG, Carline JD, Blank LL, Wenrich MD. Feasibility of hospital-based use of peer ratings to evaluate the performances of practicing physicians. Acad Med. 1996;71(4):364-70.

25. Goodyear HM, Lakshminarayana I, Wall D, Bindal T. A multisource feedback tool to assess ward round leadership skills of senior paediatric trainees: (2) Testing reliability and practicability. Postgrad Med J. 2015;91(1075):268-73.

26. Health Professions Act, R.S.A. 2000, c H-7

27. Lockyer JM, Violato C, Fidler HM. What multisource feedback factors influence physician self-assessments? A five-year longitudinal study. Acad Med. 2007;82(10 Suppl):S77-80.

28. Frank J, Snell L, Sherbino J. CanMEDS 2015 Physician Competency Framework. Ottawa: Royal College of Physicians and Surgeons of Canada; 2015.

29. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs. 2008;62(1):107-15.

30. Wood L, Hassell A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design. Med Teach. 2006;28(7):e185-91.

31. Kain NA, Hodwitz K, Yen W, Ashworth N. Experiential knowledge of risk and support factors for physician performance in Canada: a qualitative study. BMJ Open. 2019 Feb 22;9(2)

32. Cunningham CT, Quan H, Hemmelgarn B, Noseworthy T, Beck CA, Dixon E, et al. Exploring physician specialist response rates to web-based surveys. BMC Med Res Methodol. 2015;15(1):32.

33. Tambor ES, Chase GA, Faden RR, Geller G, Hofman KJ, Holtzman NA. Improving response rates through incentive and follow-up: the effect on a survey of physicians' knowledge of genetics. Am J Public Health. 2011;83(11):1599-603.

34. Choung RS, Locke GR, Schleck CD, Ziegenfuss JY, Beebe TJ, Zinsmeister AR, et al. A low response rate does not necessarily indicate non-response bias in gastroenterology survey research: a population-based study. J Public Health. 2012;21(1):87-95.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

35. Curtin R, Presser S, Singer E. The effects of response rate changes on the index of consumer sentiment. Public Opinion Quarterly. 2000;64(4):413-28.

36. Keeter S, Kennedy C, Dimock M. Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. Public Opinion Quarterly. 2006;70:759-79.

37. Visser PS, Krosnick JA, Marquette J. Mail surveys for election forecasting? An evaluation of the columbus dispatch poll. Public Opinion Quarterly. 1996;60:181-227.

38. Decieux JP, Mergener A, Neufang KM, Sischka P. Implementation of the forced answering option within online surveys: Do higher item response rates come at the expense of participation and answer quality? Psihologija. 2015;48(4):311-26.

39. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Public Opinion Quarterly. 2006;5(70):646-75.

40. Nohr EA, Frydenberg M, Henriksen TB, Olsen J. Does low participation in cohort studies induce bias? Epidemiology (Cambridge, Mass). 2006;17(4):413-8.

41. Holbrook A, Krosnick J, Pfent A. The causes and consequences of response rates in surveys by the news media and government contractor survey research firms. In: Lepkowski JM, editor. Advances in telephone survey methodology. New York: Wiley; 2007. p. 1-92.

42. Smither JW, London M, Reilly RR. Does Performance Improve Following Multisource Feedback? a Theoretical Model, Meta-Analysis, and Review of Empirical Findings. Personnel Psychology. 2005;58(1):33-66.

43. Lockyer J, Violato C, Fidler H. Likelihood of change: a study assessing surgeon use of multisource feedback data. Teach Learn Med. 2003;15(3):168-74.

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

STROBE Statement—Checklist of items that should be included in reports of ***cross-sectional studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
	page 1,2	(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	p 5,6 2	Explain the scientific background and rationale for the investigation being reported
Objectives	p 7 3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	p 7 4	Present key elements of study design early in the paper
Setting	p 7 5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	p 7 6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants
Variables	p 8 7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	p 7,8 8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	p 7,8 9	Describe any efforts to address potential sources of bias
Study size	p 7 10	Explain how the study size was arrived at
Quantitative variables	p 8 11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	p 8 12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, describe analytical methods taking account of sampling strategy
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
	p 9 Table 1	(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	p 9 14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
	Table 1	(b) Indicate number of participants with missing data for each variable of interest
Outcome data	p 9,10 15*	Report numbers of outcome events or summary measures
Main results	p 9,10 16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
	Table 2,3	(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses

For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml

BMJ Open: first published as 10.1136/bmjopen-2020-037610 on 19 July 2020. Downloaded from http://bmjopen.bmj.com/ on April 9, 2024 by guest. Protected by copyright.

Discussion			
Key results	p10	18	Summarise key results with reference to study objectives
Limitations	p11	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	p10-12	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	p13	21	Discuss the generalisability (external validity) of the study results
Other information			
Funding	p17	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

BMJ Open: first published as 10.1136/bmjopen-2020-037610 on 19 July 2020. Downloaded from http://bmjopen.bmj.com/ on April 9, 2024 by guest. Protected by copyright.