

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 (<u>http://bmjopen.bmj.com</u>).

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

Are weak or negative clinical recommendations associated with higher geographic variation in utilization than strong or positive recommendations? Cross-sectional study of 24 health care services

Journal:	BMJ Open
Manuscript ID	bmjopen-2020-044090
Article Type:	Original research
Date Submitted by the Author:	24-Aug-2020
Complete List of Authors:	Ulyte, Agne; University of Zurich, EBPI Wei, Wenjia; University of Zurich, EBPI Gruebner, Oliver; University of Zurich, EBPI Bähler, Caroline ; Helsana Group, Department of Health Sciences Brüngger, Beat; Helsana Group, Department of Health Sciences Blozik, Eva; Helsana Group, Department of Health Sciences von Wyl, Viktor ; University of Zurich, EBPI Schwenkglenks, M; University of Zurich Dressel, Holger; University of Zurich
Keywords:	EPIDEMIOLOGY, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, HEALTH SERVICES ADMINISTRATION & MANAGEMENT
	ADMINISTRATION & MANAGEMENT

SCHOLARONE[™] Manuscripts



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 <u>licence</u>.

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 <u>Creative Commons</u> 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.

relievont

2	
2 3 <u>1</u> 4 5	Are weak or negative clinical recommendations associated with higher geographic variation in
6 7 2 8	utilization than strong or positive recommendations? Cross-sectional study of 24 health care
9 10 3 11 12	services
13 4 14	
¹⁵ 5 16	Authors
¹⁷ 18 19	Agne Ulyte ^a *, Wenjia Wei ^a , Oliver Gruebner ^{a,b} , Caroline Bähler ^{a,c} , Beat Brüngger ^{a,c} , Eva Blozik ^{c,d} , Viktor von Wyl ^a ,
20 7 21	Matthias Schwenkglenks ^a , Holger Dressel ^e
22 8 23	
²⁴ 9	Affiliations
26 2710	a - Department of Epidemiology, Epidemiology, Biostatistics & Prevention Institute, University of Zurich,
28 2911 30	Hirschengraben 84, 8001 Zurich, Switzerland
31 <mark>12</mark> 32	b – Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
³³ 13	Hospital Zurich, Hirschengraben 84, 8001 Zurich, Switzerland
35 36 <mark>14</mark> 37	c - Department of Health Sciences, Helsana Group, Zürichstrasse 130, 8600 Dübendorf, Switzerland
3815 39	d - Institute of Primary Care, University of Zurich and University Hospital Zurich, Pestalozzistrasse 24, 8091
40 <u>1</u> 6 41	Zürich, Switzerland
⁴² 17 43	e – Division of Occupational and Environmental Medicine, Department of Epidemiology, Epidemiology,
44 4518	Biostatistics & Prevention Institute, University of Zurich and University Hospital Zurich, Hirschengraben 84, 8001
46 4719 48 4920 50	Zurich, Switzerland
⁵¹ 21 52	*corresponding author
⁵³ 54 22	Agne Ulyte
55 56 23 57	Email: agne.ulyte@uzh.ch
58 2 4 59 60	Phone: +41 44 63 44681

1	
² 25	Address: University of Zurich, Epidemiology, Biostatistics and Prevention Institute (EBPI), Hirschengraben 84,
4	8001 Zurich, Switzerland
₅ 26	SUCE ZUIICH, SWITZEHAID
6	
7	
8	
9	
10	
11	
12	
13	
14 15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29 30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40 41	
41 42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54 55	
55 56	
50 57	
58	
59	
60	

(

1	
2 3 27 4	Abstract
5 6 28 7	Objectives
7 8 29 9	When research evidence is lacking, patient and provider preferences, expected to vary geographically, might
1030 11	have a stronger role in clinical decisions. In this study, we investigated whether the strength or the direction of
¹² 31 13	recommendation is associated with the degree of geographic variation in utilization.
14 1532 16	Design
1733 18	In this cross-sectional study, we selected 24 services following a comprehensive approach. The strength and
1934 20	direction of recommendations were assessed in duplicate. Multilevel models were used to adjust for
²¹ 35 22	demographic and clinical characteristics and estimate unwarranted variation.
²³ 2436	Setting
25 26 37 27	Observational study of claims to mandatory health insurance in Switzerland in 2014.
2838 29	Participants
3039 31	Enrolees eligible for the 24 health care services.
³² 3340	Primary outcome measures
34 3541 36	The resulting variances of regional random effects, also expressed as median odds ratio (MOR). Services
374 2 38	grouped by strength and direction of recommendations were compared with Welch's t-test.
³⁹ 43 40	Results
41 42 42	The sizes of the eligible populations ranged from 1992 to 409 960 patients. MOR ranged between 1.13 for
43 4445	aspirin in secondary prevention of myocardial infarction to 1.68 for minor surgical procedures performed in
45 46 46 47	inpatient instead of outpatient settings. Services with weak recommendations had a negligibly higher variance
48 <mark>47</mark> 49	and MOR (difference in means [95CI%] 0.03 [-0.06, 0.11] and 0.05 [-0.11, 0.21]). Services with negative
⁵⁰ 48 51	recommendations had a slightly higher variance and MOR (difference in means [95Cl%] 0.07 [-0.03, 0.18] and
52 5349	0.14 [-0.06, 0.34]).
54 5550 56	Conclusions
50 5751 58	In this exploratory study, the geographic variation in the utilization of services associated with strong versus
⁵⁹ 52 60	weak recommendations was not substantially different. The geographic variation of services associated with
60	weak recommendations was not substantially anterent. The geographic variation of services associated with

BMJ Open

² ₃ 54	between the strength and direction of recommendations and the variation may be indirect or modified by other
4 5 55 6	characteristics of services. As initiatives discouraging low-value care are gaining attention worldwide, these
7 56 8 9 57 10	findings may inform future research in this area.
¹¹ 58 12	Strengths and limitations of this study
13 1459 15	Although the strength and direction of recommendations is generally expected to influence the
1660 17	variation in clinical decisions, this is the first study to analyse this relationship quantitatively.
¹⁸ 61 19	• The effect of the strength and direction of a recommendation on the geographic variation in health care
²⁰ 62 21	utilization was assessed within a comprehensive set of 24 health care services.
22 2363	Unwarranted variation of the services utilization was extracted with a single standard approach.
24 2564 26	 Indirect relationship and modifiers of the effect could not be studied.
²⁷ 65 28	
²⁹ 30 ⁶⁶	Keywords: geographic variation in health care; unwarranted variation; clinical recommendations; clinical
31 32 67	practice guidelines; evidence-based medicine; low-value care.
33 34 25	
35 36 37	
38 39	
40 41	
42 43	
44 45	
46 47	
48 49	
50 51	
52 53	
54 55	
56	
57 58	
59	
60	

¢

Background

1

According to the evidence-based medicine (EBM) framework, clinical decisions should be guided by research evidence, clinical circumstances, and patient preferences, and be integrated with clinical expertise [1]. If evidence is weak or lacking, patient preferences and clinical expertise have a particularly strong role in the decision [2, 3]. In a clinical practice guideline, such a situation would be reflected by a weak recommendation [2]. As patient preferences tend to vary geographically [4], and physician practice styles are also significantly influenced by the region of practice [5, 6], clinical decisions associated with less conclusive research evidence or weak recommendations may have higher geographic variation.

Surprisingly, there is little direct evidence whether weak recommendations are, in fact, associated with higher variation. The few available studies focused on a single specialty and did not quantify the variation in a uniform way, complicating the comparison of results [7, 8]. Therefore, despite many studies highlighting the substantial geographic variation in the utilization of various health care services [9–11], it is not clear, what role the components of the EBM framework play for this variation.

A second potential contributor to different services having different degrees of geographic variation is the direction of recommendation: positive (prescriptive) or negative (proscriptive), as in Choosing Wisely recommendations [12]. Negative recommendations usually concern long-used low-value practices and are based on the lack of supporting evidence or evidence of harms [12–14]. In contrast to positive recommendations, which often introduce new services or indications, negative recommendations usually challenge existing practices that are justified primarily by clinical expertise and judgement, expected to vary regionally [4]. Positive and negative recommendations have different perceived barriers to their implementation [15], which could contribute to different variation patterns as well. However, no study has directly compared the geographic variation associated with positive and negative recommendations.

The primary aim of this study was to assess whether health care services with weak recommendations are associated with higher geographic variation in utilization. In addition, a secondary aim was to test the association of geographic variation with the direction of the recommendation.

1 2	
3 95 4	Methods
5 6 96	Study hypotheses
7 8 97 9	Although this is primarily an explorative study, we formulated two specific hypotheses. The primary hypothesis
1098 11	of the study was that health care services with weaker evidence, as reflected in weak recommendations in
¹² 99 13	clinical guidelines, would have higher geographic variation in utilization than those with strong
14 1500	recommendations. The secondary hypothesis was that services with negative (proscriptive) recommendations
16 11701 18 11202	would have higher geographic variation compared to those with positive (prescriptive) recommendations.
20 2103 22	Selection of studied health care services
22 23 2104 24	This study was part of a project assessing the geographic variation of the utilization of a set of health care
24 25 <u>2</u> 1 0 5	services in Switzerland [16]. Studied health care services were translated from selected recommendation
27 21806	statements in clinical practice guidelines, following a systematic approach. We collected clinical practice
29 3 407 31	guidelines of Swiss, European, and applicable international medical societies, used in Switzerland and guiding
31 32 3108	the care for major non-communicable diseases (as defined by the Swiss Federal Office of Public Health, FOPH
33 34 3 5 09	[17]). Recommendation statements from selected clinical practice guidelines were considered pragmatically by
36 31710	the authors according to their clinical relevance, the expected frequency of service use, and the size of the
38 3911 40	eligible population. Identified recommended or discouraged services were then screened for feasibility of
40 41 42 42	measuring the utilization in eligible populations with Swiss health insurance claims data, based on an approach
42 43 4413	described earlier [18].
45 41614 47	We aimed for the selected services to reflect both strong and weak, positive and negative
4615 49	recommendations, as well as different health care services types. We focused particularly on outpatient primary
50 51 51	health care services, as they are relevant to the biggest part of the population. However, we also included some
⁵² தி17	discouraged services outside primary health care to extend the spectrum of populations investigated.
54 5 5 18	The final selection comprised 24 services, including services for screening (N=4), diagnosis (N=6),
56 51/19 58	primary prevention (N=1), treatment (N=4), and secondary prevention (N=9). Definitions of the selected
59 120 60	services are provided in Supplementary file 1.

121

2 122 ,123 7124 9125 15 11628 17 1<u>6</u>29 19 24 2632 26 21733 28 33 3436 35 31637 37 ³138 39 40 4139 4440 44 4541 46 4742 <u>5</u>1,44 53 51445 55 5646

Assessment of recommendations: strength and direction

23 Once the services were selected, their associated recommendations were formally assessed.

For each service, we selected the guideline in which the service was originally identified, and also looked up corresponding guidelines by the relevant European, American, and international clinical societies. From this set of guidelines, we selected for assessment the one that was the most applicable to Switzerland in 2014 (see Supplementary file 2 for the prioritization algorithm). We did not consider the guidelines of Swiss medical societies, as their quality of reporting is partially low [19], and they tend to be consistent with European and international guidelines. If the service was initially selected based on a guideline published after 2014, and no applicable guideline could be identified for 2014, the recommendation was automatically considered weak.

Thus, a single recommendation statement was assessed for each service. The assessment was done in duplicate by two authors (AU and HD, both medical doctors). Discordant judgements were resolved with mutual agreement in a discussion. Each recommendation was classified as strong or weak (corresponding to GRADE definition [20]), and positive or negative. The algorithm and criteria for the classification are detailed in Supplementary file 2 for the strength, and in Supplementary file 3 for the direction of a recommendation. The list of guidelines containing the recommendation statements that were assessed is provided in Supplementary file 4.

9 Swiss health insurance claims data

The utilization of the selected health care services was evaluated using mandatory health insurance claims data from the Helsana Group, covering approximately 1.2 million people (15% of the Swiss population). Helsana Group is one of several private companies providing mandatory health insurance in Switzerland. Eligible patient populations were identified from the patients enrolled with Helsana in 2014. Patients with incomplete address information, living in nursing homes and receiving reimbursement via lump-sums (masking some outpatient services), asylum seekers, those living outside Switzerland, and Helsana employees were excluded.

The data provided by Helsana were anonymized. According to the national ethical and legal regulations, The data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the national ethical and legal regulations, the data provided by Helsana were anonymized. According to the data provided by the data

BMJ Open

1 2, 10	
2149 3149 4	
4 5150 6	Models of geographic variation
7151 8	The utilization of each health care service was determined for each member of the eligible population (see
9 152 10	Supplementary file 1 for definitions of the populations and services). For each service, the resulting binary
11 153 12	outcome variable was modelled with a multilevel logistic regression technique, using 106 Swiss MobSpat regions
13 ₁ 1454	("mobilité spatiale"), as defined by the Swiss Federal Statistical Office [21], as the higher level. Each study
15 1 1655 17	participant's residence was assigned to the corresponding MobSpat region.
1 1 256 19	Fixed effects were estimated for the following explanatory variables: age, sex, number of comorbidities
²⁰ 2157 21	(0, 1, 2, and 3 or more), and clinical characteristics of relevance for specific indicators (see Supplementary file 1).
22 2358	These variables are often viewed as associated with warranted variation [22]. In contrast, we did not adjust for
24 2 1 559 26	variables associated with unwarranted variation (e.g., insurance characteristics or provider density). From each
20 21/60 28	multilevel model, we extracted the variance of the regional random effects, reflecting the potentially
²⁹ 161 30	unwarranted geographic variation. We also converted the variance to median odds ratios (MORs) for more
$^{31}_{32}_{32}$	convenient interpretation [23, 24] and plotting. MOR is interpreted as the median odds of service utilization by
33 31463 35	two individuals with identical characteristics in two randomly selected regions. As MOR is directly extrapolated
31664 37	from the variance, the ranking of these two parameters coincides.
³² 65 39	
40 166 41	Statistical analysis of the hypotheses
42 43 <mark>6</mark> 7	Variances of the regional random effects of services utilization from the models were used as data points in the
44 45568 46	final analysis of the hypotheses. Variances of services associated with weak and strong recommendations, as
4769 48	well as negative and positive recommendations were compared with Welch's unequal variances t-test. Mean
49 5070	differences and 95% confidence intervals were presented. The same analysis was also performed for MOR, to
51 5 1 5 1 71	improve interpretability of detected group differences.
53 5 1472 55	Although the number of the services analysed was rather small (24), the distribution of the analysed
55 5 <u>6</u> 73 57	variances was deemed sufficiently close to normal to warrant the use of parametric tests. To account for the
58 174 59	small and unequal sample sizes, we used Welch's t-test, which is considered more robust in this setting [25].
60 175	Confidence intervals were not adjusted for multiple testing.
	8

Statistical analyses were performed using R 3.6.0 [26] and MLwiN 3.01 [27] integrated in STATA 14.2 using the runmlwin package [28].

Patient and public involvement

This study was performed as part of the National Research Programme 74 "Smarter Health Care" of the Swiss National Science Foundations. Patients and public, including policy makers and healthcare services providers, are involved in interpreting, disseminating and translating the overall results of studies conducted under this programme. Representatives of patients, health care providers, health insurers, and health care policy makers are members of the advisory board of the project. They provided feedback on the planned study design and its preliminary results. Individual patients were not directly involved in the planning and conducting of this study.

Results

Characteristics of the eligible populations and the geographic variation of the services are shown in Table 1. Across the services, the sizes of the eligible populations ranged from 1 992 patients with a new diseasemodifying antirheumatic drug (DMARD) prescription to 409 960 patients with recommended influenza vaccination. MOR, reflecting potentially unwarranted geographic variation in utilization of the services, ranged from 1.13 [1.02-1.29] for aspirin in secondary prevention of myocardial infarction (MI) to 1.68 [1.53-1.87] for minor surgical procedures performed in inpatient instead of outpatient settings.

Category	Health care service (abbreviated)	Utilization in eligible population	Eligible popula	tion		Recomment	lation	Random effects i	n multileve
			Total N	Mean age (sd)	Female N (%)	Strength	Direction	Variance	Median o ratio (MO
Screening	Colon cancer screening	5.9%	276387	58.6 (5.8)	142675 (51.6%)	Strong		0.04[0.03-0.06]	1.21[1.17
-	Breast cancer screening	20.9%	178145	61.0 (7.2)	178145 (100%)	Weak N	Positive	0.22[0.16-0.29]	1.56[1.47
	Prostate cancer screening	28.4%	145874	59.1 (6.2)	0 (0%)	Weak	Positive	0.07[0.05-0.10]	1.29[1.25
	Osteoporosis screening	4.9%	60812	72.6 (8.7)	60812 (100%)	Weak S	Positive	0.08[0.04-0.13]	1.31[1.22
	DM: HbA1c test	69.6%	49198	66.6 (13.0)	22138 (45.0%)	Strong	Positive	0.17[0.12-0.23]	1.48[1.40
	DM: renal function test	44.3%	49198	66.6 (13.0)	22138 (45.0%)	Strong of	Positive	0.06[0.04-0.09]	1.27[1.22
	DM: LDL test	44.3%	33975	60.1 (11.2)	13977 (41.2%)	Strong 3	Positive	0.13[0.09-0.19]	1.42[1.34
Diagnosis	DM: eye examination	55.5%	49198	66.6 (13.0)	22138 (45.0%)	Weak 3	Positive	0.07[0.05-0.10]	1.29[1.24
	TSH screening	76.1%	169232	56.8 (18.5)	111847 (66.1%)	Strong	Negative	0.18[0.13-0.25]	1.50[1.42
	POCR	13.0%	47215	60.3 (17.2)	27086 (57.4%)	Strong	Negative	0.18[0.13-0.26]	1.50[1.40
Primary prevention	Influenza vaccination	20.9%	409960	64.1 (16.3)	230202 (56.2%)	Strong	Positive	0.04[0.03-0.05]	1.20[1.17
Treatment	Benzodiazepines	18.6%	243951	75.0 (7.6)	141986 (58.2%)	Strong 📑	Negative	0.14[0.10-0.18]	1.42[1.36
	Proton pump inhibitors	55.5%	153523	55.7 (17.8)	93543 (60.9%)	Weak	Negative	0.02[0.02-0.03]	1.16[1.13
	Inpatient procedures	61.4%	10656	50.5 (13.7)	7719 (72.4%)	Weak o	Negative	0.30[0.20-0.43]	1.68[1.53
	Caesarean section	28.5%	9449	31.9 (5.1)	9449 (100%)	 Weak	 Negative 	0.05[0.02-0.09]	1.24[1.16
Secondary	AMI: aspirin	47.0%	2232	72.4 (13.7)	801 (35.9%)	Strong =	Positive	0.02[0.00-0.07]	1.13[1.02
prevention	AMI: statin	34.2%	2232	72.4 (13.7)	801 (35.9%)	Strong	Positive	0.14[0.06-0.27]	1.43[1.25
	AMI: beta-blocker	42.1%	2232	72.4 (13.7)	801 (35.9%)	Strong	Positive	0.05[0.00-0.13]	1.25[1.05
	AMI: ACE/ARB	43.8%	2232	72.4 (13.7)	801 (35.9%)	Strong		0.04[0.00-0.12]	1.21[1.03
	AMI: P2Y12 inhibitors	46.8%	2232	72.4 (13.7)	801 (35.9%)	Strong Q	Positive	0.03[0.00-0.10]	1.18[1.04
	PPI with NSAID	43.5%	95072	61.0 (16.2)	60804 (64.0%)	Strong	Positive	0.02[0.01-0.03]	1.15[1.12
	PAD: statin	28.5%	23868	63.6 (16.5)	12113 (50.7%)	Strong	Positive	0.04[0.03-0.07]	1.22[1.17
	Afib: anticoagulation	27.5%	8291	80.8 (7.9)	4037 (48.7%)	Strong 6	Positive	0.05[0.02-0.09]	1.24[1.16
	GCC with new DMARD	58.7%	1992	59.2 (15.3)	1369 (68.7%)	Strong to Weak	Positive	0.06[0.01-0.18]	1.27[1.07

Page 12 of 33

		BMJ Open BMJ Open P	age 12 of
1	196	SD – standard deviation, N – number, DM – diabetes mellitus, LDL – low density lipid, TSH – thyroid-stimulating hormone, POCR – performance chest radiography, AMI – acute	
2 3	197	myocardial infarction, ACE/ARB – angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers, PPI – proton pump is hibitors, NSAID – nonsteroidal anti-inflamma	tory
4 5	198	drugs, PAD – peripheral artery disease, Afib – atrial fibrillation, GCC – glucocorticosteroid drugs, DMARD – disease-modifying antirhequmatic drug.	
6 7	199		
8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41		10 May 2021. Downleaded from http://binicoen.tami.com/ on April 17, 2024 by guest. Protected by copyright	
42 43 44		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	11

BMJ Open

2 2 3 200	For three services, a major guideline relevant in 2014 in Switzerland could not be identified (long-term
4 ₅201	use of proton pump inhibitors, minor inpatient surgery procedures, elective Caesarean section). A total of eight
6 7202 8	services had weak, and six services had negative underlying recommendations. Median MOR was 1.29 for
9203 10	services with weak and 1.25 for services with strong recommendations; 1.26 for services with positive and 1.46
11 12 12	for services with negative recommendations (Figure 1).
13 1 <mark>24</mark> 05 15	
1 2: 06 17	Figure 1 Geographic variation of the health care services grouped by strength (A) and direction (B) of
1 2807 19	recommendations
20 208 21	A Weak and strong recommendations; B Positive and negative recommendations. MOR – median odds ratio.
22 2309 24	Boxplots depict the interquartile range of values (upper and lower hinges), and the median value.
22510	
26 22/11 28	Based on Welch's t-test, the difference in mean variances [95Cl%] of services with weak and strong
²⁰ 2212 30	recommendations was 0.03 [-0.06, 0.11], and the difference in mean MOR was 0.05 [-0.11, 0.21]. The difference
$31 \\ 32 \\ 32 \\ 32 \\ 32 \\ 32 \\ 32 \\ 32 \\ $	in mean variances [95CI%] of services with negative and positive recommendations was 0.07 [-0.03, 0.18], and
33 32414 35	the difference in mean MOR was 0.14 [-0.06, 0.34].
33 32414 35 32615 37	
33 3414 35 3615 37 38 3916	
33 32414 35 32615 37	the difference in mean MOR was 0.14 [-0.06, 0.34].
33 3214 35 3515 37 3916 40 41 4217 43 4218	the difference in mean MOR was 0.14 [-0.06, 0.34]. Discussion
33 32414 35 32615 37 38 3916 40 41 42 42 43 42 43 44 45 4619	the difference in mean MOR was 0.14 [-0.06, 0.34]. Discussion We did not find a direct association between the strength of clinical recommendation and the geographic
33 3214 35 3515 37 3916 40 41 42 42 43 4218 45	the difference in mean MOR was 0.14 [-0.06, 0.34]. Discussion We did not find a direct association between the strength of clinical recommendation and the geographic variation in the utilization of 24 health care services. The geographic variation in the utilization of services with
33 3414 35 3615 37 39 40 41 42 43 441 43 441 45 4619 47	the difference in mean MOR was 0.14 [-0.06, 0.34]. Discussion We did not find a direct association between the strength of clinical recommendation and the geographic variation in the utilization of 24 health care services. The geographic variation in the utilization of services with underlying negative recommendations was slightly higher than for those with positive recommendations. In
33 3414 35 3615 37 38 40 41 42 42 42 43 45 45 45 45 47 45 47 45 47 45 47 45 45 47 45 45 47 5021 52 57 52 57 222	the difference in mean MOR was 0.14 [-0.06, 0.34]. Discussion We did not find a direct association between the strength of clinical recommendation and the geographic variation in the utilization of 24 health care services. The geographic variation in the utilization of services with underlying negative recommendations was slightly higher than for those with positive recommendations. In general, moderate potentially unwarranted geographic variation was observed, with MOR smaller than 1.50 for
33 3414 35 3615 37 38 40 41 42 42 43 44 45 4619 47 45 47 4920 5221 5222 54 5223	the difference in mean MOR was 0.14 [-0.06, 0.34]. Discussion We did not find a direct association between the strength of clinical recommendation and the geographic variation in the utilization of 24 health care services. The geographic variation in the utilization of services with underlying negative recommendations was slightly higher than for those with positive recommendations. In general, moderate potentially unwarranted geographic variation was observed, with MOR smaller than 1.50 for all but one service.
33 3414 35 3615 37 38 40 41 42 42 42 43 45 4619 47 45 47 45 47 4920 5021 522 512 52 54 5223 56 5224 55 52 54 55 55 52 54 55 55 52 55 52 55 52 55 52 55 52 58	the difference in mean MOR was 0.14 [-0.06, 0.34]. Discussion We did not find a direct association between the strength of clinical recommendation and the geographic variation in the utilization of 24 health care services. The geographic variation in the utilization of services with underlying negative recommendations was slightly higher than for those with positive recommendations. In general, moderate potentially unwarranted geographic variation was observed, with MOR smaller than 1.50 for all but one service. At least two other studies have to some extent examined the association between the strength of
33 3414 35 3615 37 38 40 41 42 43 45 47 4217 43 45 47 4218 47 45 47 420 5221 52 52 54 5223 56 57 52 54 5223 56 57 52 54 55 52 54 55 52 54 55 52 54 55 52 54 55 52 54 55 52 54 55 52 54 55 52 54 55 52 54 55 52 54 55 54 55 52 54 55 52 54 55 52 54 55 52 54 55 52 54 55 54 52 54 55 54 52 54 55 54 52 54 55 54 52 54 55 54 52 54 55 52 54 55 54 52 54 55 54 52 54 55 54 52 54 55 54 52 54 55 54 52 54 55 54 52 54 55 54 52 54 55 54 52 54 56 57 52 54 56 57 57 54 56 57 54 57 54 56 57 74 56 57 74 56 57 74 56 57 74 56 57 74 56 57 74 7	the difference in mean MOR was 0.14 [-0.06, 0.34]. Discussion We did not find a direct association between the strength of clinical recommendation and the geographic variation in the utilization of 24 health care services. The geographic variation in the utilization of services with underlying negative recommendations was slightly higher than for those with positive recommendations. In general, moderate potentially unwarranted geographic variation was observed, with MOR smaller than 1.50 for all but one service. At least two other studies have to some extent examined the association between the strength of recommendations and the variation in adherence, each focusing on a single clinical specialty. In et al [7],

found that surgeon practices in knee and hip arthroplasty in Australia varied regardless of the strength of
evidence available. Potentially, different clinical areas could be associated with different barriers to guideline
implementation, modifying the relationship between recommendations and variation.

To better understand why a direct association of recommendation strength and variation in adherence was not observed, it may be useful to revisit the EBM framework [1]. The EBM framework is *normative* and defines how clinical decisions *should* be made [1, 29]. However, this may not always coincide with how decisions *are* made – a process analysed by *descriptive* theories [29]. In fact, the EBM model has been developed as a conceptual rather than practical guidance of evidence implementation [1], and has not yet generated a coherent theory of clinical decision making, and in particular, of how evidence is incorporated [30]. Thus, although a direct relationship between the strength of recommendation and the geographic variation of service utilization would be encouraged by the normative EBM framework, it may not always be observed.

There are numerous reasons why even strong recommendations [31, 32], or conclusive research evidence more broadly [33], may not directly translate into clinical practice. Research on knowledge translation has identified multiple barriers at different levels of the health care system, including structural, organizational, peer-group, and professional factors [34] – many of which depend on the specific context where a service is provided, and thus may vary geographically. Knowledge transfer processes are highly non-linear and rely on triggering mechanisms [35]. Factors external to research evidence significantly affect translation – potentially creating large geographic heterogeneity even within services with strong recommendations.

An influential framework, explaining different degrees of variation between health care services, has been proposed by Wennberg and colleagues [36]. According to this framework, services are classified into effective, preference-sensitive, and supply-sensitive care. Effective care (services based on solid evidence, so that virtually all patients would choose them) largely corresponds to services with strong recommendations, as defined by GRADE and applied in this study [2]. Preference-sensitive care partly corresponds to services with weak recommendations, as they both imply trade-offs of risks and benefits of multiple options of care [36]. The utilization of preference-sensitive surgical procedures usually has higher variation than of those associated with effective care [37]. In contrast, supply-sensitive care defines the frequency, setting, and intensity of care provision rather than specific types of health care services. It is associated with high, supply-related variation,

54 but is rarely discussed in guidelines [37], and therefore, could not be included in our study. However, the service 55 of minor surgeries performed as inpatient instead of outpatient procedures could be considered close to the 56 supply-sensitive category. In fact, it had the highest MOR (1.68) in our study.

Regarding the secondary hypothesis, we found that services associated with negative recommendations had slightly higher geographic variation. We did not find other studies directly comparing the regional variation of services with the direction of recommendations. Few studies, focusing mostly on low-value care, have reported MOR as an expression of geographic variation, further limiting the comparison. For example, in a study by Badgery-Parker et al [38], services discouraged by *Choosing Wisely* were shown to have regional MOR from 1.1 to 2.6 – a range that includes all of our observed MORs.

Negative recommendations usually address a widespread service that lacks supporting evidence of benefit or the benefit is outweighed by harms [2]. In contrast to services with positive recommendations, which are introduced after supporting evidence is produced, services with negative recommendations typically become part of the clinical practice *before* evidence is sufficient to rule out their overall benefit. Therefore, their use could be related more to clinical expertise and practice, and could be expected to vary locally. Indeed, the barriers to implementing positive and negative recommendations seem to be different [15] – signalling that the pathways how they are interpreted and integrated into clinical decisions might also be different. As Choosing Wisely and similar initiatives are increasingly gaining attention [39], our finding of higher geographic variation associated with negative recommendations may inform future research and implementation strategies.

This exploratory study has several limitations. First, although we aimed at a balanced selection of clinical fields and service types, the number (24) and range of studied services was limited by the data source. Swiss claims data lack information on outpatient diagnoses, inpatient treatment details, and clinical information such as test results [18]. Lack of clinical information also meant that some populations were not as specific as defined by the recommendation. For example, beta-blockers and angiotensin-converting-enzyme (ACE) inhibitors are recommended after a myocardial infarction contingent on heart failure and left ventricular dysfunction. As such clinical details were unavailable, we had to rely on them being present in the majority of the hospitalized myocardial infarction cases and distributed equally geographically. However, we believe that estimates of variation are accurate, as each of the 24 data points was generated by multilevel modelling of utilization in

Page 16 of 33

BMJ Open

populations of 85000 patients on average, including all major explanatory variables such as age, sex, and indicators of morbidity. Second, the services studied were unavoidably different by characteristics other than the strength or direction of recommendation, such as service type or clinical area, potentially resulting in confounding. Third, both the observed utilization and its geographic variation depend on the definition of the service and population [40]. We aimed to measure the *unwarranted* variation in utilization by using servicespecific denominators (eligible populations) and adjusting for relevant clinical characteristics. How exactly *unwarranted* and *warranted* variation should be defined and measured, and what adjustments are necessary to differentiate them, is debated [22, 41]. Fourth, the grouping of recommendations by strength and direction was partly subjective, although we tried to make it reproducible with a clear algorithm, implemented in duplicate. Unfortunately, many different systems for evaluating the strength of recommendations exist [42], which cannot be easily reconciled, and the most prominent, GRADE approach, is not always explicitly used.

To explore the studied questions further, the sample of services could be expanded to inpatient and specialist care. Further, a meta-study of the numerous individual studies of geographic variation in health care services could be undertaken. However, this would currently be challenging, as studies choose different adjustment variables and specificity of studied populations, and report the variation in different quantitative forms (e.g., MOR, systematic component of variation, range). Furthermore, there is a need for qualitative studies of the reasons for the variability of clinical decisions, and how clinical expertise in these decisions interacts with evidence, clinical circumstances, and patient preferences. Qualitative evidence could help to generate more complex hypotheses for further quantitative studies, built on a finer understanding of all the factors influencing the variability of clinical decisions. Specialty-specific sets of services could also be further investigated.

Conclusions

In this exploratory study of 24 health care services mostly in the outpatient primary care setting, we did not observe a significant difference in the degree of geographic variation in utilization of services associated with strong versus weak recommendations. Services associated with negative recommendations had slightly higher geographic variation. The relationship between the strength of recommendations and the variation may be

BMJ Open

indirect or significantly modified by other characteristics of services, such as service type or clinical area. As
 initiatives discouraging low-value care are gaining attention worldwide, these findings may inform future
 research in this area.

12 Funding

This work was supported by the Swiss National Science Foundation (SNSF) National Research Program "Smarter
 Health Care" (NRP 74), as part of project number 26, grant number 407440_167349.

6 **Competing interests**

MS declares a grant from Helsana Insurance Group, outside the submitted work. Helsana Group provided support in the form of salaries for authors BB, EB and CB, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The other authors declare no competing interests.

2 Authors' contributions

MS, VvW and HD developed the underlying study program. AU and HD developed study design, with support from all authors. AU, WW, CB, EB, BB did data extraction, preparation and management. WW, MS and OG developed multilevel models applied in this study. AU and WW analysed the data. AU drafted the manuscript, with major inputs from HD and MS, and contributions from all authors. All authors read and approved the final manuscript.

29 Data sharing statement

The data underlying this study cannot be shared publicly because they are the property of Helsana (https://www.helsana.ch/en/helsana-group), and have restricted public access on grounds of patient privacy. The data are managed by Helsana and subsets of the database are available for researchers after request and under specific conditions. Data are available from Helsana (gesundheitskompetenz@helsana.ch) for researchers who meet the criteria for access to confidential data. Helsana will consider the possibilities of the research

proposal and decide to grant access if the research questions can be answered with use of the Helsana data. 336 When requests are granted, data are accessible only in a secure environment. 7337 8 1338 References 11 1. Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient 1340 choice. Evid Based Med. 2002;7:36-8. 16 13741 2. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from 18 ¹342 evidence to recommendations: The significance and presentation of recommendations. J Clin Epidemiol. 2013;66:719–25. doi:10.1016/j.jclinepi.2012.03.013. 23 **⊰**44 3. Andrews JC, Schünemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going 25 2645 from evidence to recommendation - Determinants of a recommendation's direction and strength. J Clin 27 2846 Epidemiol. 2013;66:726–35. doi:10.1016/j.jclinepi.2013.02.003. 29 4. Cutler D, Skinner JS, Stern AD, Wennberg D. Physician beliefs and patient preferences: A new look at regional variation in health care spending. Am Econ J Econ Policy. 2019;11:192–221. 34 33549 5. Finkelstein A, Gentzkow M, Williams H. Sources of Geographic Variation in Health Care: Evidence From 36 3750 PatientMigration. Q J Econ. 2016;131:1681–726. 38 6. Molitor D. The evolution of physician practice styles: Evidence from cardiologist migration. Am Econ J Econ Policy. 2018;10:326-56. 43 43453 7. In H, Neville BA, Lipsitz SR, Corso KA, Weeks JC, Greenberg CC. The role of National Cancer Institute-45 43554 designated cancer center status: observed variation in surgical care depends on the level of evidence. Ann Surg. 2012;255:890-5. doi:10.1097/SLA.0b013e31824deae6. 8. Mayer M, Naylor J, Harris I, Badge H, Adie S, Mills K, et al. Evidence base and practice variation in acute care processes for knee and hip arthroplasty surgeries. PLoS One. 2017;12:e0180090. \$357 54 53558 doi:10.1371/journal.pone.0180090. 56 ⁵⁷59 9. Corallo AN, Croxford R, Goodman DC, Bryan EL, Srivastava D, Stukel TA. A systematic review of medical practice variation in OECD countries. Health Policy (New York). 2014;114:5–14. 361 doi:10.1016/j.healthpol.2013.08.002.

2362 3 ³ 62	10. The Dartmouth Atlas of Health Care. https://www.dartmouthatlas.org/. Accessed 11 May 2020.
4 5363	11. Public Health England. Public Health Profiles. https://fingertips.phe.org.uk/profile/atlas-of-variation.
6 7364	Accessed 11 May 2020.
8 9 365 10	12. Cassel CK, Guest JA. Choosing wisely: Helping physicians and patients make smart decisions about their care.
11 366 12	JAMA - Journal of the American Medical Association. 2012;307:1801–2.
13 1 <mark>3</mark> 67	13. Elshaug AG, Watt AM, Mundy L, Willis CD. Over 150 potentially low-value health care practices: an Australian
15 13668 17	study. Med J Aust. 2012;197:556–60.
1 3 69 19	14. Prasad V, Vandross A, Toomey C, Cheung M, Rho J, Quinn S, et al. A decade of reversal: An analysis of 146
²⁰ 370 21	contradicted medical practices. Mayo Clin Proc. 2013;88:790–8.
22 371 23	15. Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: A meta-synthesis of GPs' attitudes to clinical
24 2 3572	practice guidelines. Br J Gen Pract. 2007;57:971–8.
26 2 373 28	16. National Research Programme 74. Project 26: How do guidelines and recommendations influence medical
²⁹ 374 30	treatment? http://www.nfp74.ch/en/projects/healthcare-across-sectors/project-schwenkglenks. Accessed 15
31 3275	Apr 2020.
33 33476	17. Swiss Federal Office of Public Health and Swiss Conference of Cantonal Health Directors. Nationale Strategie
35 3 3577 37	Prävention nichtübertragbarer Krankheiten (NCD-Strategie) 2017-2024 [National Strategy of Non-communicable
³³⁸ 78 39	Diseases]. 2016.
40 4179	18. Ulyte A, Bähler C, Schwenkglenks M, von Wyl V, Gruebner O, Wei W, et al. Measuring diabetes guideline
42 4 3 80 44	adherence with claims data: systematic construction of indicators and related challenges. BMJ Open.
44 4 5 81 46	2019;9:e027138. doi:10.1136/bmjopen-2018-027138.
4 3 82 48	19. Bachmann L, Ulyte A, Dressel H. Clinical practice guidelines of medical societies in Switzerland: analysis of
49 383 50	the current state. Swiss Med Wkly. 2019;149:w20134.
51 5 3 84	20. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. GRADE: Going from Evidence to
53 5 3 485 55	Recommendations. BMJ. 2008;336:1048–51.
5 586 57	21. Bundesamt für Statistik. MS-Regionen. https://www.bfs.admin.ch/bfs/de/home/statistiken/raum-
58 387 59	umwelt/nomenklaturen/msreg.assetdetail.415729.html. Accessed 11 May 2020.
⁶⁰ 388	22. Sutherland K, Levesque J. Unwarranted clinical variation in health care: Definitions and proposal of an

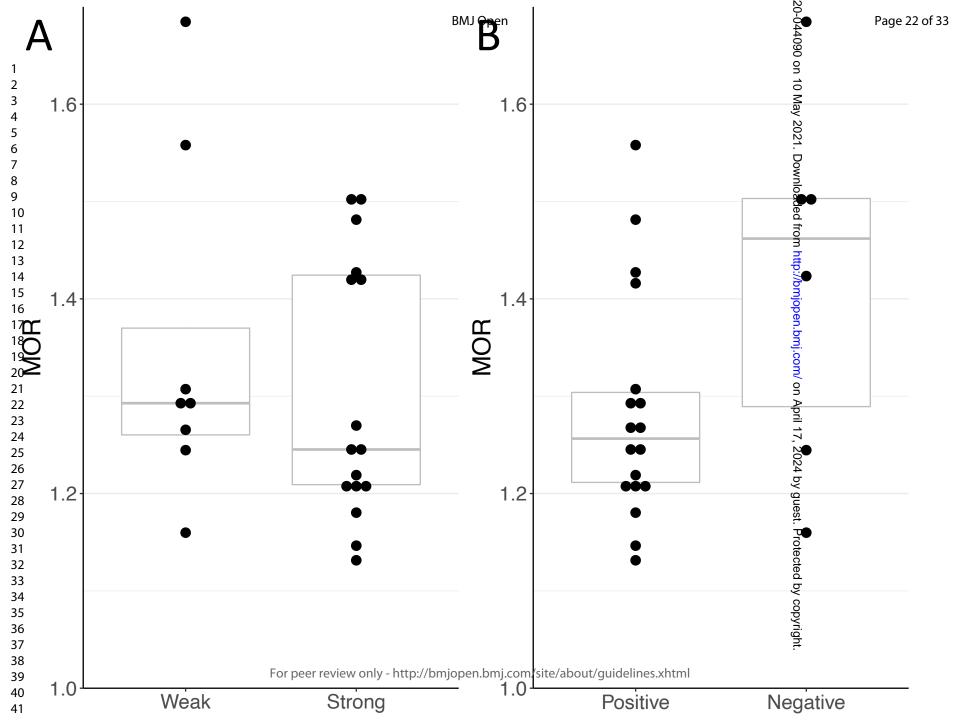
¢

1	
2 389 3	analytic framework. J Eval Clin Pract. 2019;:jep.13181. doi:10.1111/jep.13181.
4 5390	23. Larsen K, Petersen JH, Budtz-Jørgensen E, Endahl L. Interpreting parameters in the logistic regression model
6 7391 8	with random effects. Biometrics. 2000;56:909–14.
9 392 10	24. Larsen K, Merlo J. Appropriate Assessment of Neighborhood Effects on Individual Health: Integrating
11 393 12	Random and Fixed Effects in Multilevel Logistic Regression. Am J Epidemiol. 2005;161:81–8.
13 1394	25. Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-
15 1 36 95 17	Whitney U test. Behav Ecol. 2006;17:688–90.
1 3896 19	26. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical
²⁰ 397 21	Computing, Vienna, Austria. 2019. https://www.r-project.org/. Accessed 26 Nov 2019.
22 398 23	27. Charlton, C, Rasbash, J, Browne, WJ, Healy, M, Cameron B. MLwiN Version 3.01. Centre for Multilevel
24 2 3 99 26	Modelling, University of Bristol. 2017.
24700 28	28. Leckie G, Charlton C. Runmlwin: A program to run the MLwiN multilevel modeling software from within
2401 30	Stata. J Stat Softw. 2013;52:1–40.
$^{31}_{32}_{32}$	29. Djulbegovic B, Elqayam S, Dale W. Rational decision making in medicine: Implications for overuse and
33 3403 35	underuse. J Eval Clin Pract. 2018;24:655–65. doi:10.1111/jep.12851.
34604 37	30. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. The Lancet.
³⁴⁰⁵ 39	2017;390:415–23.
40 4106	31. Francke AL, Smit MC, De Veer AJE, Mistiaen P. Factors influencing the implementation of clinical guidelines
42 4 5 07 44	for health care professionals: A systematic meta-review. BMC Med Inform Decis Mak. 2008;8:38.
44 4508 46	doi:10.1186/1472-6947-8-38.
4709 48	32. Lugtenberg M, Zegers-Van Schaick JM, Westert GP, Burgers JS. Why don't physicians adhere to guideline
49 410 50	recommendations in practice? An analysis of barriers among Dutch general practitioners. Implement Sci.
$51 \\ 5411$	2009;4:1–9.
53 54412 55	33. Wallace J, Nwosu B, Clarke M. Barriers to the uptake of evidence from systematic reviews and meta-
55 54913 57	analyses: A systematic review of decision makers' perceptions. BMJ Open. 2012;2:e001220.
58 414 59	34. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement
60 415	Sci. 2012;7:50. doi:10.1186/1748-5908-7-50.

Page 21 of 33

BMJ Open

1	
² 3416	35. Braithwaite J, Churruca K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation
4 5417 6	science: A theoretical and empirical analysis of systems change. BMC Med. 2018;16:63. doi:10.1186/s12916-
7418 8	018-1057-z.
9 419 10	36. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over medicare reform. Health Affairs. 2003;22
11 420 12	SUPPL.
13 421	37. Wennberg JE. Tracking medicine : a researcher's quest to understand health care. Oxford University Press;
15 1 4622 17	2010.
1 4623 19	38. Badgery-Parker T, Feng Y, Pearson S-A, Levesque J-F, Dunn S, Elshaug AG. Exploring variation in low-value
20 21 21	care: a multilevel modelling study. BMC Health Serv Res. 2019;19:345.
22 425	39. Levinson W, Kallewaard M, Bhatia RS, Wolfson D, Shortt S, Kerr E a. "Choosing Wisely": a growing
24 2 4526 26	international campaign. BMJ Qual Saf. 2015;24:167–74. doi:10.1136/bmjqs-2014-003821.
2 4727 28	40. Schwartz AL, Landon BE, Elshaug AG, Chernew ME, McWilliams JM. Measuring low-value care in Medicare.
²⁹ 30	JAMA Intern Med. 2014;174:1067–76. doi:10.1001/jamainternmed.2014.1541.
31 429 33	41. Harrison R, Manias E, Mears S, Heslop D, Hinchcliff R, Hay L. Addressing unwarranted clinical variation: A
34430 35	rapid review of current evidence. J Eval Clin Pract. 2019;25:53–65. doi:10.1111/jep.12930.
34631 37	42. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. Systems for grading the quality of evidence
³⁴⁸ 32 39	and the strength of recommendations I: Critical appraisal of existing approaches. BMC Health Serv Res.
40 41 41 33	2004;4:1–7.
42 4 3 34	2004;4:1-7.
44 45 435 46	Supplementary files
47	
48 49 50	Supplementary file 1 Definitions and descriptions of the studied health care services and eligible populations
50 51 57	Supplementary file 2 Algorithm and criteria for the assessment of the strength of recommendation
52 ∯38 54	Supplementary file 3 Algorithm and criteria for the assessment of the direction of recommendation
54539 56	Supplementary file 4 List of guidelines selected for the study, describing the services analysed



1	
2 3	
4	
5 6	
6 7	
8	
9	
10 11	
12	
13	
14 15	
16	
10 11 12 13 14 15 16 17	
18 19	
20	
21	
22 23	
24	
25	
26 27	
28	
29	
30 31	
31 32	
33	
34 35	
36	
37	
38 39	
40	
41	
42 43	
45 44	
45	
46	

Additional file 1 Definitions and descriptions of the studied health care services and eligible population	าร
--	----

daltional	file 1 Definitions a	and descriptions of the studied	health care services a	nd eligible populations 0,000 000 000 000 000 000 000 000 000	
Category	Health care	Service description and	Eligible population	Recommendation 4000	Specific clinical explanatory
	service	frequency		0 on	variables
creening	Colon cancer	Colonoscopy/ year	Anyone 50-69 years old	Colonoscopy should be done ever \vec{e}_{10}	Previous treatment of cancer o
	screening			years for people 50-69 years old.	inflammatory bowel disease,
				2021.	hospitalization with colon
				Dow	disease in the last year
	Breast cancer	Mammography/year	50-74 years old women	Mammography should be done every 2	Previous treatment of breast or
	screening			years for 50-74 years old women. 👸	other cancer
	Prostate cancer	Prostate-specific antigen (PSA)	50-70 years old men	Early detection of prostate cancer	Previous treatment of cancer,
	screening	testing/ year		(opportunistic screening) should be	hospitalization with prostate
				offered to the well-informed man	disease in the last year
	Osteoporosis	Dual-energy x-ray	Women over 60 and	DXA densitometry is recommended for	Presence of more than one risk
	screening	absorptiometry (DXA)/ year	with risk factors ^a of	postmenopausal women with	factor
			spontaneous fractures	spontaneous fractures or increase	
				of them.	
Diagnosis	DM: HbA1c test	Glycated haemoglobin (HbA1c)	Adult drug-treated	HbA1c test should be done for diagetes	Oral diabetes medication or
		test twice/ year	diabetes patients	patients at least twice a year. $\frac{1}{3}$	insulin
	DM: renal	Albuminuria and serum	Adult drug-treated	Albuminuria and serum creatinine ests	Oral diabetes medication or
	function test	creatinine tests/ year	diabetes patients	should be done for diabetes patie	insulin
				least once a year.	
	DM: LDL test	Low-density lipoprotein (LDL)	Adult drug-treated	LDL test should be done for diabet es	Oral diabetes medication or
		test/ year	diabetes patients under	patients at least once a year.	insulin
			75 years old	patients at least once a year. fected by copyright.	

3 4

			BMJ Open	6/bmjopen-	
	DM: eye	Ophthalmologist visit/ year	Adult drug-treated	Eye exam should be performed for	Oral diabetes medication or
	examination		diabetes patients	diabetes patients at least once a year.	insulin
	TSH screening	Thyroid-stimulating hormone	Adults without thyroid	TSH should be measured as an inital	-
		(TSH) test without T3 and T4	disease and receiving	screening test for d	
		tests on the same day	TSH test	hypo/hyperthyroidism, while T3 and T4	
				test should follow if TSH is abnorn	
	POCR	Outpatient preoperative chest	Adult patients with	Routine chest radiography is not	-
		radiography (POCR) up to 2	inpatient surgical	recommended before surgery.	
		months before surgery	procedures	Routine chest radiography is not recommended before surgery.	
Primary	Influenza	Influenza outpatient vaccination/	People over 65 years	People over 65 years old and paties	Hospitalization with pneumonia
prevention	vaccination	year	old or with a specified	with chronic conditions, specified by	in the last year
			chronic condition ^b	Federal Office of Public Health, should	
				be vaccinated against influenza every	
			e e	year.	
Freatment	Benzodiazepines	Cumulative prescription of	Anyone over 65 years	Long-term use of benzodiazepinesand	Treated epilepsy, stay in a
		benzodiazepines (BZD) for >8	old	other hypnotics is discouraged for old	nursing home in the last year,
		weeks/ year		patients.	hospitalization in the last year
				patients.	with a diagnosis indicative of
				, 202	justified benzodiazepine use
	Proton pump	Cumulative prescription of	Adults receiving PPI or	PPI should not be used at maximal dose	-
	inhibitors	proton pump inhibitors (PPI) or	H2 drugs	for prolonged periods of time.	
		H2 histamine receptor			
		antagonists (H2) for >8 weeks/		rotec	
		year		Protected by copyright.	
	1	1	1	co	1
				ppyrię	
				ht.	

Page 25 of 33

33			BMJ Open	6/bmjopen-2	
	T			0	1
	Inpatient	Specified surgical procedures ^c	Adult patients with	If none of the special conditions apply,	-
	procedures	done in the outpatient setting	specified surgical	certain surgical procedures should be	
			procedures (either as	done in the outpatient setting.	
			in- or outpatient)	n 10	
	Caesarean	Caesarean section (C-section)	Women giving birth	C-section should not be performe	-
	section		without absolute	unless absolute or relative indications	
			indications ^d for C-	are present.	
			section	are present.	
Secondary	AMI: aspirin	Aspirin prescription within 2	Adult patients with AMI	All myocardial infarction patients aould	Hospitalization for stroke or
prevention		weeks after acute myocardial		take aspirin long-term. ਰ੍ਹੇ	bleeding event or prescribed
		infarction (AMI)	-0r	take aspirin long-term. from http://b	anticoagulation in the last yea
	AMI: statin	High-dose statin prescription	Adult patients with AMI	All myocardial infarction patients around	Hospitalization for stroke in th
		within 2 weeks after AMI		get statins long-term.	last year
				get statins long-term.	
	AMI: beta-	Beta-blocker prescription within	Adult patients with AMI	All myocardial infarction patients with	Hospitalization with heart
	blocker	2 weeks after AMI		heart failure or impaired function \underline{P}	failure diagnosis in the last yea
				should get beta-blockers long-term	
	AMI: ACE/ARB	Angiotensin-converting enzyme	Adult patients with AMI	All myocardial infarction patients with	-
		(ACE) or angiotensin receptor		heart failure or impaired function $\frac{24}{5}$	
		blocker (ARB) antihypertensive		should get ACE or ARB antihypertensive	
		medication prescription within 2		medication long-term.	
		weeks after AMI			
	1	1		<u>ୁ</u> ଜୁ ପ	

by copyright.

		BMJ Open	6/bmjopen	
AMI: P2Y12	P2Y12 antiplatelet drug ^e	Adult patients with AMI	All myocardial infarction patients abould	Hospitalization for a bleeding
inhibitors	prescription within 2 weeks after		get P2Y12 antiplatelet drugs for atleast	event or prescribed
	AMI		1-12 months according to the bleeding	anticoagulation in the last year
			risk profile and AMI treatment.	
PPI with NSA	ID PPI prescription within 1 month	Adult patients with a	Patients taking long-term NSAID and	Concurrent use of antiplatelet,
	or up to 3 months before initial	cumulative NSAID	with risk factors for gastric ulcer ^f spould	anticoagulation drugs or oral
	long-term nonsteroidal anti-	prescription of >8	also take PPI.	glucocorticoids, hospitalization
	inflammatory drug (NSAID)	weeks at maximal dose		for bleeding event in the last
	prescription		also take PPI.	year.
PAD: statin	Prescription of statins within 3	Adult patients		-
	months after peripheral artery	undergoing diagnostic	patients with PAD.	
	disease (PAD) identification	or treatment	Statins are recommended for all from http://bmjo	
		procedures for PAD	njo po	
Afib:	Oral anticoagulation prescription	Adult patients with	All patients with atrial fibrillation sould	-
anticoagulat	on within 2 weeks after atrial	atrial fibrillation	be prescribed oral anticoagulation for	
	fibrillation (Afib) identification	diagnosis and	embolic events prevention according to	
		additional risk factors ^g	the CHA ₂ DS ₂ -VASc score.	
GCC with ne	W Glucocorticoid (GCC) prescription	Adult patients with a	Short-term glucocorticoids should be	-
DMARD	within 1 month or up to 3	new prescription of	taken with newly prescribed DMA BD.	
	months before disease-	DMARD by a	24 by guest. P	
	modifying antirheumatic drug	rheumatologist	/ gue	
	(DMARD) prescription		St. F	
		1		1]

a. Recent distal radius, proximal humerus, vertebral or femoral fracture, use of drugs increasing the risk of osteoporosis, use of oral glacocorticoids, diabetes, ankylosing spondylitis, osteogenesis imperfecta, rheumatoid arthritis, inflammatory bowel disease, Cushing's disease, alcohol or nicotine abuse, Aronic liver disease, gastrectomy, malnutrition, hypogonadism, hyper- or hypothyroidism, and hyperparathyroidism. Patients currently treated or diagnosed with osteo prosis were excluded. /right.

BMJ Open

6/bmjopen-2

90 on 10 May

- b. Cardiovascular disease, chronic pulmonary disease, diabetes, chronic liver disease, renal failure, immune deficiency, systemic neuro Bgic disorders.
- c. Varicose veins ligation and stripping, surgical procedures of haemorrhoids, inguinal hernia and cervix, knee arthroscopy and menisco tomy, tonsillectomy.
- d. Placental, umbilical cord or fetal pathology, HIV or genital HSV infection, or multiple pregnancy.
- e. Clopidogrel, prasugrel or ticagrelor.
- f. Concurrent use of antiplatelet, anticoagulant drugs, oral glucocorticoids or recent hospitalization with any major bleeding.
- g. Risk factors (congestive heart failure, hypertension, age 65-74 or \geq 75 years old, diabetes, previous stroke, transient ischemic attack, or thromboembolism, cardiovascular disease, female sex) were extracted from available claims data and summed according to CHA2DS2-VASc score. Patients with CHA2DS $\frac{1}{2}$ VASc score of \geq 2 for males and \geq 3 for females were included.
- DM diabetes mellitus, HbA1c Glycated haemoglobin, LDL low density lipid, TSH thyroid-stimulating hormone, T3 and T4 triiod with vronine and thyroxine. POCR preoperative chest radiography, BZD – benzodiazepines, PPI – proton pump inhibitors, H2 – H2 histamine receptor antagonists, C-section – Caesarean section, AMI – acute myocardial infarction, ACE/ARB – angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers, NSAID – nonsteroidal anti-inflammatory drugs, PAD – peripheral artery disease, Afib – atrial fibrillation, GCC – glucocorticosteroid drugs, DMARD – disease-modifying antirheumatic drug. //bmjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright Pen.uny...

N of authors	Steps	
Single	1.	Identify the relevant medical societies for each selected health care service.
	2.	Look up the European or international medical societies' websites and journals, and
		identify relevant guidelines published before 2014. In addition, look up if Swiss federal
		legislation guidelines exist by 2014.
	3.	If none found, look up American medical society and identify relevant guidelines
		published before 2014.
	4.	If none found, consider the recommendation weak.
In duplicate	5.	Once the guideline and the recommendation statement are located, classify the
		recommendation into strong or weak ^a .
	-	Strong recommendation implies that the desirable effects of adherence to a
		recommendation outweigh the undesirable effects.
	-	That means that most informed patients would choose the recommended
		management and that clinicians can structure their interactions with patients
		accordingly.
	-	For clinicians, that would mean that most patients should receive the recommended
		course of action.
		For patients, that would mean that most people in such a situation would want the
		recommended course of action and only a small proportion would not; patients should
		request discussion if the intervention is not offered.
	-	Weak recommendation implies that the desirable effects of adherence to a
		recommendation probably outweigh the undesirable effects, but the guideline panel i less confident.
	-	Thus, a weak recommendation is conditional or optional, and means that patients'
		choices will vary according to their values and preferences, and clinicians must ensure
		that patients' care is in keeping with their values and preferences.
	-	For clinicians, that would mean that they should recognize that different choices will b
		appropriate for different patients and that they must help each patient to arrive at a
		management decision consistent with her or his values and preferences.
		For patients, that would mean that most people in such situation would want the
		recommended course of action, but many would not.
a adapted from	: Guvatt G	GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. GRADE: Going from Eviden
	-	ЛJ 2008;336:1048–51.

Additional file 2 Algorithm and criteria for the assessment of the strength of recommendation

N of authors
In duplicate

Recommendation	Reference	Comment
Colon cancer screening	[1]	
Breast cancer screening	[2]	
Prostate cancer screening	[3]	
Osteoporosis screening	[4]	
DM: HbA1c test	[5]	
DM: renal function test	[5]	
DM: LDL test	[5]	
DM: eye examination	[5]	
TSH screening	[6,7]	
POCR	[8]	
Influenza vaccination	[9]	
Benzodiazepines	[10]	
Proton pump inhibitors		Swiss national guideline since 2016 [11]
Inpatient procedures	- 0	Swiss federal regulation exists from 2019 [12]
Caesarean section	-	Swiss national guideline since 2015 [13]
AMI: aspirin	[14,15]	
AMI: statin	[14,15]	
AMI: beta-blocker	[14,15]	
AMI: ACE/ARB	[14,15]	
AMI: P2Y12 inhibitors	[14,15]	
PPI with NSAID	[16]	
PAD: statin	[17]	
Afib: anticoagulation	[18]	
GCC with new DMARD	[19]	

Additional file 4 List of guidelines selected for the study, describing the services analysed

DM – diabetes mellitus, HbA1c – Glycated haemoglobin, LDL – low density lipid, TSH – thyroid-stimulating hormone, POCR – preoperative chest radiography, AMI – acute myocardial infarction, ACE/ARB – angiotensinconverting enzyme inhibitors or angiotensin II receptor blockers, PPI – proton pump inhibitors, NSAID – nonsteroidal anti-inflammatory drugs, PAD – peripheral artery disease, Afib – atrial fibrillation, GCC – glucocorticosteroid drugs, DMARD – disease-modifying antirheumatic drug.

- Lansdorp-Vogelaar I, Karsa L. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition – Introduction. Endoscopy. 2012 Sep 25;44(S 03):SE15–30.
- Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med. 2009 Nov 17;151(10):716.
- Heidenreich A, Bellmunt J, Bolla M, Joniau S, Mason M, Matveev V, et al. EAU Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Treatment of Clinically Localised Disease. Eur Urol. 2011 Jan;59(1):61–71.

Page 31 of 33

BMJ Open

1 2	4.	Kanis JA, McCloskey E V., Johansson H, Cooper C, Rizzoli R, Reginster J-Y. European guidance for the
3		diagnosis and management of osteoporosis in postmenopausal women. Osteoporos Int. 2013 Jan
4		19;24(1):23–57.
5 6	5.	Rydén L, Grant PJ, Anker SD, Berne C, Cosentino F, Danchin N, et al. ESC Guidelines on diabetes, pre-
7		diabetes, and cardiovascular diseases developed in collaboration with the EASD. Eur Heart J. 2013 Oct
8 9		14;34(39):3035–87.
10	c	
11	6.	Bahn RS, Burch HB, Cooper DS, Garber JR, Greenlee MC, Klein I, et al. Hyperthyroidism and Other Causes of
12 13		Thyrotoxicosis: Management Guidelines of the American Thyroid Association and American Association of
14		Clinical Endocrinologists. Thyroid. 2011 Jun;21(6):593–646.
15 16	7.	Garber J, Cobin R, Gharib H, Hennessey J, Klein I, Mechanick J, et al. Clinical Practice Guidelines for
17		Hypothyroidism in Adults: Cosponsored by the American Association of Clinical Endocrinologists and the
18		American Thyroid Association. Endocr Pract. 2012 Nov;18(6):988–1028.
19 20	8.	Choosing Wisely. American College of Surgeons. Admission pre-op chest x-rays. September 4, 2013.
21		Available from: https://www.choosingwisely.org/clinician-lists/american-college-surgeons-admission-or-
22 23		preop-chest-x-ray-on-ambulatory-patients/ Accessed on May 11, 2020
24	9.	Swiss Federal Office of Public Health. Recommendations for Influenza Vaccination (Empfehlungen zur
25	5.	Grippeimpfung). 2011.
26 27	4.0	
28	10.	Choosing Wisely. American Geriatrics Society. Benzodiazepines sedative hypnotics for insomnia in older
29 30		adults. February 21, 2013. Available from: https://www.choosingwisely.org/clinician-lists/american-
31		geriatrics-society-benzodiazepines-sedative-hypnotics-for-insomnia-in-older-adults/ Accessed on May 11,
32		2020
33 34	11.	Swiss Society of General Internal Medicine. Smarter Medicine. Top-5 List for Ambulatory Care. 2016.
35		Available from: https://www.smartermedicine.ch/de/top-5-listen/ambulante-allgemeine-innere-
36 37		medizin.html Accessed on May 11, 2020
38	12.	Swiss Federal Office of Public Health. "Outpatient instead of Inpatient" [Änderung der Krankenpflege-
39		Leistungsverordnung (KLV) betreffend «Ambulant vor Stationär»]. Available from:
40 41		https://www.bag.admin.ch/bag/de/home/versicherungen/krankenversicherung/krankenversicherung-
42		
43 44		revisionsprojekte/konsultation-ambulant-vor-stationaer.html Accessed on May 11, 2020
45	13.	Hoesli I, Alma-Stucki S El, Drack G, Girard T, Irion O, Schulzke S, et al. Guideline Sectio Caesarea. 2015;1–20.
46		Available from: https://www.sggg.ch/fachthemen/guidelines/ Accessed on May 11, 2020
47 48	14.	Hamm CW, Bassand J-P, Agewall S, Bax J, Boersma E, Bueno H, et al. ESC Guidelines for the management of
49		acute coronary syndromes in patients presenting without persistent ST-segment elevation: The Task Force
50 51		for the management of acute coronary syndromes (ACS) in patients presenting without persistent ST-
52		segment elevatio. Eur Heart J. 2011 Dec 1;32(23):2999–3054.
53	15.	Steg PG, James SK, Atar D, Badano LP, Lundqvist CB, Borger MA, et al. ESC Guidelines for the management
54 55		of acute myocardial infarction in patients presenting with ST-segment elevation. Eur Heart J. 2012 Oct
56		1;33(20):2569–619.
57	16	
58 59	16.	Recommendations for use of selective and nonselective nonsteroidal antiinflammatory drugs: An American
60		College of Rheumatology white paper. Arthritis Rheum. 2008 Aug 15;59(8):1058–73.

- 17. Tendera M, Aboyans V, Bartelink M-L, Baumgartner I, Clement D, Collet J-P, et al. ESC Guidelines on the diagnosis and treatment of peripheral artery diseases: Document covering atherosclerotic disease of extracranial carotid and vertebral, mesenteric, renal, upper and lower extremity arteries * The Task Force on the Diagnosis and Treat. Eur Heart J. 2011 Nov 2;32(22):2851–906.
- 18. Camm AJ, Lip GYH, De Caterina R, Savelieva I, Atar D, Hohnloser SH, et al. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation. Eur Heart J. 2012 Nov 1;33(21):2719-47.
- 19. Smolen JS, Landewé R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis. 2014 Mar;73(3):492–509.

<text>

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

Page 33 of 33

		BMJ Open	
	STI	ROBE 2007 (v4) Statement—Checklist of items that should be included in reports of <i>cross-sectional studies</i>	
Section/Topic	Item #	Recommendation 9	Reported on page
Title and abstract	1	(<i>a</i>) Indicate the study's design with a commonly used term in the title or the abstract	1,2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction	•	221.	
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4
Objectives	3	State specific objectives, including any prespecified hypotheses	4,5
Methods			
Study design	4	Present key elements of study design early in the paper	2,
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-7
Participants	6	(<i>a</i>) Give the eligibility criteria, and the sources and methods of selection of participants	7, Supplementary file 1
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7, Supplementary file 1
Data sources/	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe	6, Supplementary
measurement		comparability of assessment methods if there is more than one group	files 1-4
Bias	9	Describe any efforts to address potential sources of bias	6,7
Study size	10	Explain how the study size was arrived at	5
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which grougings were chosen and why	7
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	7
		(b) Describe any methods used to examine subgroups and interactions	7
		(b) Describe any methods used to examine subgroups and interactions \overline{c} (c) Explain how missing data were addressed \overline{c}	NA
		(d) If applicable, describe analytical methods taking account of sampling strategy	NA
		(e) Describe any sensitivity analyses So Image: Solution of the sensitivity analyses So	NA
Results			

 6/bmjopen-2

Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examine of individuals at each stage of study—eg numbers potentially eligible, examine of the stage of study and the study and th	8,9
		confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9
		(b) Indicate number of participants with missing data for each variable of interest	NA
Outcome data	15*	Report numbers of outcome events or summary measures	9
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision $\frac{4}{3}$ eg, 95% confidence	9, Supplementary
		interval). Make clear which confounders were adjusted for and why they were included	file 1
		(b) Report category boundaries when continuous variables were categorized	NA
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13,14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14,15
Generalisability	21	Discuss the generalisability (external validity) of the study results	14
Other information			
Funding	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	15

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in controls in case-control studies.

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 are http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

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

Are weak or negative clinical recommendations associated with higher geographic variation in utilization than strong or positive recommendations? Cross-sectional study of 24 health care services

Journal:	BMJ Open
Manuscript ID	bmjopen-2020-044090.R1
Article Type:	Original research
Date Submitted by the Author:	19-Feb-2021
Complete List of Authors:	Ulyte, Agne; University of Zurich, EBPI Wei, Wenjia; University of Zurich, EBPI Gruebner, Oliver; University of Zurich, EBPI Bähler, Caroline ; Helsana Group, Department of Health Sciences Brüngger, Beat; Helsana Group, Department of Health Sciences Blozik, Eva; Helsana Group, Department of Health Sciences von Wyl, Viktor ; University of Zurich, EBPI Schwenkglenks, M; University of Zurich Dressel, Holger; University of Zurich
Primary Subject Heading :	Health services research
Secondary Subject Heading:	Evidence based practice, Health policy, Patient-centred medicine
Keywords:	EPIDEMIOLOGY, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Protocols & guidelines < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, HEALTH SERVICES ADMINISTRATION & MANAGEMENT
	·

SCHOLARONE[™] Manuscripts



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 <u>licence</u>.

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 <u>Creative Commons</u> 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.

R. O.

2	
3 <u>1</u> 4 5	Are weak or negative clinical recommendations associated with higher geographic variation in
6 7 2 8	utilization than strong or positive recommendations? Cross-sectional study of 24 health care
9 10 3 11 12	services
13 4 14	
¹⁵ 5 16	Authors
¹⁷ 18 6	Agne Ulyteª*, Wenjia Weiª, Oliver Gruebnerª, ^b , Caroline Bählerª,c, Beat Brünggerª,c, Eva Blozik ^{c,d} , Viktor von Wylª,
19 20 7	Matthias Schwenkglenks ^a , Holger Dressel ^e
21 22 8 23	
24 25 25	Affiliations
26 2710 28	a - Department of Epidemiology, Epidemiology, Biostatistics & Prevention Institute, University of Zurich,
28 2911 30	Hirschengraben 84, 8001 Zurich, Switzerland
31 <mark>12</mark> 32	b – Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
³³ 13 34	Hospital Zurich, Hirschengraben 84, 8001 Zurich, Switzerland
35 36 ¹ 4	c - Department of Health Sciences, Helsana Group, Zürichstrasse 130, 8600 Dübendorf, Switzerland
37 3815 39	d - Institute of Primary Care, University of Zurich and University Hospital Zurich, Pestalozzistrasse 24, 8091
40 <u>1</u> 6 41	Zürich, Switzerland
⁴² 17 43	e – Division of Occupational and Environmental Medicine, Department of Epidemiology, Epidemiology,
⁴⁴ 45 ¹⁸	Biostatistics & Prevention Institute, University of Zurich and University Hospital Zurich, Hirschengraben 84, 8001
46 4719 48 49 20 50	Zurich, Switzerland
⁵¹ 21 52	*corresponding author
⁵³ 54 22	Agne Ulyte
55 56 23 57	Email: agne.ulyte@uzh.ch
5824 59 60	Phone: +41 44 63 44681
-	

1	
² 25	Address: University of Zurich, Epidemiology, Biostatistics and Prevention Institute (EBPI), Hirschengraben 84,
4	8001 Zurich, Switzerland
5 26 6	8001 zunch, Switzenand
7	
8	
9	
10 11	
12	
13	
14	
15 16	
17	
18	
19 20	
20 21	
22	
23	
24 25	
26	
27	
28 29	
30	
31	
32 33	
34	
35	
36 37	
38	
39	
40 41	
41	
43	
44 45	
45 46	
47	
48	
49 50	
51	
52	
53 54	
55	
56	
57 58	
50 59	
60	

(

1	
2 3 27 4	Abstract
5 6 28 7	Objectives
7 8 2 9 9	When research evidence is lacking, patient and provider preferences, expected to vary geographically, might
10 30 11	have a stronger role in clinical decisions. We investigated whether the strength or the direction of
¹² 31 13	recommendation is associated with the degree of geographic variation in utilization.
14 1532 16	Design
1 <i>7</i> 33 18	In this cross-sectional study, we selected 24 services following a comprehensive approach. The strength and
19 3 4 20	direction of recommendations were assessed in duplicate. Multilevel models were used to adjust for
²¹ 35 22	demographic and clinical characteristics and estimate unwarranted variation.
23 2436 25	Setting
26 37 27	Observational study of claims to mandatory health insurance in Switzerland in 2014.
28 38 29	Participants
³⁰ 39 31	Enrolees eligible for the 24 health care services.
22	
³² 3340 34	Primary outcome measures
³² 3340 34 3541 36	Primary outcome measures The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by
34 3 <i>5</i> 41	
34 3541 36 3742 38 ³⁹ 43 40	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by
34 3541 36 3742 38 3943 40 41 42	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test.
34 3541 36 3742 38 3943 40 41 42 43 44 45	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test. Results
34 3541 36 3742 38 3943 40 41 42	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test. Results The sizes of the eligible populations ranged from 1992 to 409 960 patients. MOR ranged between 1.13 for
34 3541 36 3742 38 3943 40 41 42 43 42 43 44 45 45 4646 47 4847 49	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test. Results The sizes of the eligible populations ranged from 1992 to 409 960 patients. MOR ranged between 1.13 for aspirin in secondary prevention of myocardial infarction to 1.68 for minor surgical procedures performed in
34 3541 36 3742 38 3943 40 41 44 43 44 45 45 4646 47 4847 49 5048 51	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test. Results The sizes of the eligible populations ranged from 1992 to 409 960 patients. MOR ranged between 1.13 for aspirin in secondary prevention of myocardial infarction to 1.68 for minor surgical procedures performed in inpatient instead of outpatient settings. Services with weak recommendations had a negligibly higher variance
34 3541 36 3742 38 3943 40 41 42 43 44 45 45 4646 47 4847 49 50 48 51 52 53 49	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test. Results The sizes of the eligible populations ranged from 1992 to 409 960 patients. MOR ranged between 1.13 for aspirin in secondary prevention of myocardial infarction to 1.68 for minor surgical procedures performed in inpatient instead of outpatient settings. Services with weak recommendations had a negligibly higher variance and MOR (difference in means [95Cl%] 0.03 [-0.06, 0.11] and 0.05 [-0.11, 0.21]) compared to strong
34 3541 36 3742 38 3943 40 41 44 43 44 45 45 4646 47 4847 49 5048 51	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test. Results The sizes of the eligible populations ranged from 1992 to 409 960 patients. MOR ranged between 1.13 for aspirin in secondary prevention of myocardial infarction to 1.68 for minor surgical procedures performed in inpatient instead of outpatient settings. Services with weak recommendations had a negligibly higher variance and MOR (difference in means [95CI%] 0.03 [-0.06, 0.11] and 0.05 [-0.11, 0.21]) compared to strong recommendations. Services with negative recommendations had a slightly higher variance and MOR (difference in means [95CI%] 0.03 [-0.06, 0.11] and 0.05 [-0.11, 0.21]) compared to strong
34 3541 36 3742 38 3943 40 41 44 43 44 45 45 4646 47 4847 49 5048 51 52 5349 54 5550 56 5751 58	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test. Results The sizes of the eligible populations ranged from 1992 to 409 960 patients. MOR ranged between 1.13 for aspirin in secondary prevention of myocardial infarction to 1.68 for minor surgical procedures performed in inpatient instead of outpatient settings. Services with weak recommendations had a negligibly higher variance and MOR (difference in means [95CI%] 0.03 [-0.06, 0.11] and 0.05 [-0.11, 0.21]) compared to strong recommendations. Services with negative recommendations had a slightly higher variance and MOR (difference in means [95CI%] 0.014 [-0.06, 0.34]) compared to positive recommendations.
34 3541 36 3742 38 3943 40 41 42 43 44 45 45 4646 47 4847 49 50 48 51 52 53 49 51 52 53 49 54 55 50 56 57 51	The variances of regional random effects, also expressed as median odds ratio (MOR). Services grouped by strength and direction of recommendations were compared with Welch's t-test. Results The sizes of the eligible populations ranged from 1992 to 409 960 patients. MOR ranged between 1.13 for aspirin in secondary prevention of myocardial infarction to 1.68 for minor surgical procedures performed in inpatient instead of outpatient settings. Services with weak recommendations had a negligibly higher variance and MOR (difference in means [95CI%] 0.03 [-0.06, 0.11] and 0.05 [-0.11, 0.21]) compared to strong recommendations. Services with negative recommendations had a slightly higher variance and MOR (difference in means [95CI%] 0.014 [-0.06, 0.34]) compared to positive recommendations.

BMJ Open

² ₃ 54	direction of recommendations and the variation may be indirect or modified by other characteristics of services.
4 5 55 6	As initiatives discouraging low-value care are gaining attention worldwide, these findings may inform future
7 56 8 9 57 10	research in this area.
¹¹ 58 12	Strengths and limitations of this study
13 1459 15	Although the strength and direction of recommendations is generally expected to influence the
1660 17	variation in clinical decisions, this is the first study to analyse this relationship quantitatively.
¹⁸ 61 19	• The effect of the strength and direction of a recommendation on the geographic variation in health care
²⁰ 2162	utilization was assessed within a comprehensive set of 24 health care services.
22 2363	Unwarranted variation of the services utilization was extracted with a single standard approach.
24 2564 26	Indirect relationship and modifiers of the effect could not be studied.
²⁷ 65 28	
²⁹ 30 ⁶⁶	Keywords: geographic variation in health care; unwarranted variation; clinical recommendations; clinical
31 3 <u>2</u> 67	practice guidelines; evidence-based medicine; low-value care.
33 34 35	
36 37	
38 39	
40 41	
42 43	
44 45	
46 47	
48 49	
50 51	
52	
53 54	
55	
56 57	
57 58	
59	
60	

¢

9 Background

1

According to the evidence-based medicine (EBM) framework, clinical decisions should be guided by research evidence, clinical circumstances, and patient preferences, and be integrated with clinical expertise [1]. If evidence is weak or lacking, patient preferences and clinical expertise have a particularly strong role in the decision [2, 3]. In a clinical practice guideline, such a situation would be reflected by a weak recommendation [2]. As patient preferences tend to vary geographically [4], and physician practice styles are also significantly influenced by the region of practice [5, 6], clinical decisions associated with less conclusive research evidence or weak recommendations may have higher geographic variation.

Surprisingly, there is little direct evidence whether weak recommendations are, in fact, associated with higher variation. The few available studies focused on a single specialty and did not quantify the variation in a uniform way, complicating the comparison of results [7, 8]. Therefore, despite many studies highlighting the substantial geographic variation in the utilization of various health care services [9–11], it is not clear, what role the components of the EBM framework play for this variation.

A second potential contributor to different services having different degrees of geographic variation is the direction of recommendation: positive (prescriptive) or negative (proscriptive), as in Choosing Wisely recommendations [12]. Negative recommendations usually concern long-used low-value practices and are based on the lack of supporting evidence or evidence of harms [12–14]. In contrast to positive recommendations, which often introduce new services or indications, negative recommendations usually challenge existing practices that are justified primarily by clinical expertise and judgement, expected to vary regionally [4]. Positive and negative recommendations have different perceived barriers to their implementation [15], which could contribute to different variation patterns as well. However, no study has directly compared the geographic variation associated with positive and negative recommendations.

The primary aim of this study was to assess whether health care services with weak recommendations are associated with higher geographic variation in utilization. In addition, a secondary aim was to test the association of geographic variation with the direction of the recommendation.

1 2 3 95 4	Methods
5 6 96	Study hypotheses
7 8 97 9	Although this is primarily an explorative study, we formulated two specific hypotheses. The primary hypothesis
1098 11	of the study was that health care services with weaker evidence, as reflected in weak recommendations in
¹² 99 13	clinical guidelines, would have higher geographic variation in utilization than those with strong
14 1500	recommendations. The secondary hypothesis was that services with negative (proscriptive) recommendations
16 1 1 201 18 1 <u>1</u> 202 20	would have higher geographic variation compared to those with positive (prescriptive) recommendations.
²¹ 103 22	Selection of studied health care services
23 104 24	This study was part of a project assessing the geographic variation of the utilization of a set of health care
25 21a05	services in Switzerland [16]. Studied health care services were translated from selected recommendation
27 2 <u>1</u> 806 29	statements in clinical practice guidelines, following a systematic approach. We collected clinical practice
³¹ 07 31	guidelines of Swiss, European, and applicable international medical societies, used in Switzerland and guiding
32 3108 33	the care for major non-communicable diseases (as defined by the Swiss Federal Office of Public Health, FOPH
34 3 1 509 36	[17]). Recommendation statements from selected clinical practice guidelines were considered pragmatically by
31710 318	the authors according to their clinical relevance, the expected frequency of service use, and the size of the
³ 1211 40	eligible population. Identified recommended or discouraged services were then screened for feasibility of
$^{41}_{42}_{42}$	measuring the utilization in eligible populations with Swiss health insurance claims data, based on an approach
43 4413 45	described earlier [18].
45 4614 47	We aimed for the selected services to reflect both strong and weak, positive and negative
4β15 49	recommendations, as well as different health care services types. We focused particularly on outpatient primary
50 51 51	health care services, as they are relevant to the biggest part of the population. However, we also included some
52 5317	discouraged services outside primary health care to extend the spectrum of populations investigated.
54 5 1518 56	The final selection comprised 24 services, including services for screening (N=4), diagnosis (N=6),
50 51/19 58	primary prevention (N=1), treatment (N=4), and secondary prevention (N=9). Definitions of the selected
⁵⁹ 120 60	services are provided in Supplementary file 1.

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

2 122 ,123 7124 9125 15 11628 17 1<u>6</u>29 19 24 2632 26 21733 28 33 3436 35 31637 37 ³1938 39 40 4139 <u>4</u>340 44 4641 46 4742 <u>5</u>1,44 53 51445 55 5646 57

Assessment of recommendations: strength and direction

23 Once the services were selected, their associated recommendations were formally assessed.

For each service, we selected the guideline in which the service was originally identified, and also looked up corresponding guidelines by the relevant European, American, and international clinical societies. From this set of guidelines, we selected for assessment the one that was the most applicable to Switzerland in 2014 (see Supplementary file 2 for the prioritization algorithm). We did not consider the guidelines of Swiss medical societies, as their quality of reporting is partially low [19], and they tend to be consistent with European and international guidelines. If the service was initially selected based on a guideline published after 2014, and no applicable guideline could be identified for 2014, the recommendation was automatically considered weak.

Thus, a single recommendation statement was assessed for each service. The assessment was done in duplicate by two authors (AU and HD, both medical doctors). Discordant judgements were resolved with mutual agreement in a discussion. Each recommendation was classified as strong or weak (corresponding to GRADE definition [20]), and positive or negative. The algorithm and criteria for the classification are detailed in Supplementary file 2 for the strength, and in Supplementary file 3 for the direction of a recommendation. The list of guidelines containing the recommendation statements that were assessed is provided in Supplementary file 4.

9 Swiss health insurance claims data

The utilization of the selected health care services was evaluated using mandatory health insurance claims data from the Helsana Group, covering approximately 1.2 million people (15% of the Swiss population). Helsana Group is one of several private companies providing mandatory health insurance in Switzerland. Eligible patient populations were identified from the patients enrolled with Helsana in 2014. Patients with incomplete address information, living in nursing homes and receiving reimbursement via lump-sums (masking some outpatient services), asylum seekers, those living outside Switzerland, and Helsana employees were excluded.

The data provided by Helsana were anonymized. According to the national ethical and legal regulations, The data provided by Helsana were anonymized. According to the national ethical and legal regulations, this type of analysis. This was confirmed by a waiver of the competent ethics committee (Kantonale Ethikkommission Zürich, dated January 11, 2017, BASEC-Nr. Req-2017-00011).

48

51

BMJ Open

1	
2 3149	
4 5150	Models of geographic variation
6 7151 8	The utilization of each health care service was determined for each member of the eligible population (see
9 152 10	Supplementary file 1 for definitions of the populations and services). For each service, the resulting binary
11 153 12	outcome variable was modelled with a multilevel logistic regression technique, using 106 Swiss MobSpat regions
13 ₁ 1454 15	("mobilité spatiale"), as defined by the Swiss Federal Statistical Office [21], as the higher level. MobSpat regions
1 <u>6</u> 55 17	are constructed by combining several neighbouring municipalities based on geographic and population mobility
1 1 956 19	criteria, and are often used as intermediate-size units of analysis for scientific and regional policy purposes. Each
² μ57 21	study participant's residence was assigned to the corresponding MobSpat region.
22 2358 24	Fixed effects were estimated for the following explanatory variables: age, sex, number of comorbidities
24 2 1 559 26	(0, 1, 2, and 3 or more), and clinical characteristics of relevance for specific indicators (see Supplementary file 1).
21⁄60 28	These variables are often viewed as associated with <i>warranted</i> variation [22]. In contrast, we did not adjust for
²⁹ 161 30	variables associated with <i>unwarranted</i> variation (e.g., insurance characteristics or provider density). From each
31 <u>32</u> 62 33	multilevel model, we extracted the variance of the regional random effects, reflecting the potentially
3463 35	unwarranted geographic variation. We also converted the variance to median odds ratios (MORs) for more
3 <u>1</u> 64 37	convenient interpretation [23, 24] and plotting. MOR is interpreted as the median odds of service utilization by
³ 165 39	two individuals with identical characteristics in two randomly selected regions. As MOR is directly extrapolated
40 4166 42 4 <u>3</u> 67	from the variance, the ranking of these two parameters coincides.
44	
4 5 68 46	Statistical analysis of the hypotheses
4769 48	Variances of the regional random effects of services utilization from the models were used as data points in the

Variances of the regional random effects of services utilization from the models were used as data points in the 49 507 final analysis of the hypotheses. Variances of services associated with weak and strong recommendations, as well as negative and positive recommendations were compared with Welch's unequal variances t-test. Mean <u>5</u><u>5</u>71 53 51472 differences and 95% confidence intervals were presented. The same analysis was also performed for MOR, to 55 improve interpretability of detected group differences.

5673 57 5874 59 Although the number of the services analysed was rather small (24), the distribution of the analysed 60 175 variances was deemed sufficiently close to normal to warrant the use of parametric tests. To account for the

small and unequal sample sizes, we used Welch's t-test, which is considered more robust in this setting [25].

7 Confidence intervals were not adjusted for multiple testing.

Statistical analyses were performed using R 3.6.0 [26] and MLwiN 3.01 [27] integrated in STATA 14.2
 using the runmlwin package [28].

Patient and public involvement

This study was performed as part of the National Research Programme 74 "Smarter Health Care" of the Swiss National Science Foundations. Patients and public, including policy makers and healthcare services providers, are involved in interpreting, disseminating and translating the overall results of studies conducted under this programme. Representatives of patients, health care providers, health insurers, and health care policy makers are members of the advisory board of the project. They provided feedback on the planned study design and its preliminary results. Individual patients were not directly involved in the planning and conducting of this study.

Results

Characteristics of the eligible populations and the geographic variation of the services are shown in Table 1. Across the services, the sizes of the eligible populations ranged from 1 992 patients with a new diseasemodifying antirheumatic drug (DMARD) prescription to 409 960 patients with recommended influenza vaccination. MOR, reflecting potentially unwarranted geographic variation in utilization of the services, ranged from 1.13 [1.02-1.29] for aspirin in secondary prevention of myocardial infarction (MI) to 1.68 [1.53-1.87] for minor surgical procedures performed in inpatient instead of outpatient settings.

 6/bmjopen-2020

Table 1 Characteristics of the recommended or discouraged health care services studied

Category	Health care service	Utilization in	Eligible population			Recommendation		Random effects in multilevel m	
	(abbreviated)	eligible population							
			Total N	Mean age (sd)	Female N (%)	Strength	Direction	Variance	Median od
									ratio (MOR
Screening	Colon cancer screening	5.9%	276387	58.6 (5.8)	142675 (51.6%)	Strong Strong		0.04[0.03-0.06]	1.21[1.17-1
	Breast cancer screening	20.9%	178145	61.0 (7.2)	178145 (100%)	Weak 2021	Positive	0.22[0.16-0.29]	1.56[1.47-1
	Prostate cancer screening	28.4%	145874	59.1 (6.2)	0 (0%)	Weak .	Positive	0.07[0.05-0.10]	1.29[1.25-1
	Osteoporosis screening	4.9%	60812	72.6 (8.7)	60812 (100%)	Weak §	Positive	0.08[0.04-0.13]	1.31[1.22-1
	DM: HbA1c test	69.6%	49198	66.6 (13.0)	22138 (45.0%)	Strong of	Positive	0.17[0.12-0.23]	1.48[1.40-1
	DM: renal function test	44.3%	49198	66.6 (13.0)	22138 (45.0%)	Strong 🛱	Positive	0.06[0.04-0.09]	1.27[1.22-1
	DM: LDL test	44.3%	33975	60.1 (11.2)	13977 (41.2%)	Strong 7	Positive	0.13[0.09-0.19]	1.42[1.34-1
Diagnosis	DM: eye examination	55.5%	49198	66.6 (13.0)	22138 (45.0%)	Weak 3	Positive	0.07[0.05-0.10]	1.29[1.24-1
	TSH screening	76.1%	169232	56.8 (18.5)	111847 (66.1%)	Strong 🚦	Negative	0.18[0.13-0.25]	1.50[1.42-1
	POCR	13.0%	47215	60.3 (17.2)	27086 (57.4%)	Strong	Negative	0.18[0.13-0.26]	1.50[1.40-1
Primary	Influenza vaccination	20.9%	409960	64.1 (16.3)	230202 (56.2%)	Strong 🖁	Positive	0.04[0.03-0.05]	1.20[1.17-1
prevention						h b			
Treatment	Benzodiazepines	18.6%	243951	75.0 (7.6)	141986 (58.2%)	Strong	Negative	0.14[0.10-0.18]	1.42[1.36-1
	Proton pump inhibitors	55.5%	153523	55.7 (17.8)	93543 (60.9%)	Weak g	Negative	0.02[0.02-0.03]	1.16[1.13-1
	Inpatient procedures	61.4%	10656	50.5 (13.7)	7719 (72.4%)	Weak o	Negative	0.30[0.20-0.43]	1.68[1.53-1
	Caesarean section	28.5%	9449	31.9 (5.1)	9449 (100%)	Weak A	Negative	0.05[0.02-0.09]	1.24[1.16-1
Secondary	AMI: aspirin	47.0%	2232	72.4 (13.7)	801 (35.9%)	Strong	Positive	0.02[0.00-0.07]	1.13[1.02-1
prevention	AMI: statin	34.2%	2232	72.4 (13.7)	801 (35.9%)	Strong ,	Positive	0.14[0.06-0.27]	1.43[1.25-1
	AMI: beta-blocker	42.1%	2232	72.4 (13.7)	801 (35.9%)	Strong 2024	Positive	0.05[0.00-0.13]	1.25[1.05-1
	AMI: ACE/ARB	43.8%	2232	72.4 (13.7)	801 (35.9%)	Strong 👳	Positive	0.04[0.00-0.12]	1.21[1.03-1
	AMI: P2Y12 inhibitors	46.8%	2232	72.4 (13.7)	801 (35.9%)	Strong G	Positive	0.03[0.00-0.10]	1.18[1.04-1
	PPI with NSAID	43.5%	95072	61.0 (16.2)	60804 (64.0%)	Strong	Positive	0.02[0.01-0.03]	1.15[1.12-1
	PAD: statin	28.5%	23868	63.6 (16.5)	12113 (50.7%)	Strong	Positive	0.04[0.03-0.07]	1.22[1.17-1
	Afib: anticoagulation	27.5%	8291	80.8 (7.9)	4037 (48.7%)	Strong of	Positive	0.05[0.02-0.09]	1.24[1.16-1
	GCC with new DMARD	58.7%	1992	59.2 (15.3)	1369 (68.7%)	Weak 💆	Positive	0.06[0.01-0.18]	1.27[1.07-1

Health care services, highlighted in bold, are associated with a negative recommendation. Utilization was assessed within one year, 2014, including for services that are recommended less frequently (e.g., colon cancer screening).

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

Page 12 of 37

		BMJ Open 3 op	Page 12 of
1	199	SD – standard deviation, N – number, DM – diabetes mellitus, LDL – low density lipid, TSH – thyroid-stimulating hormone, POCR – p🗞 operative chest radiography, AMI – acute	
2 3	200	myocardial infarction, ACE/ARB – angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers, PPI – proton pump is hibitors, NSAID – nonsteroidal anti-inflamma	atory
4 5	201	drugs, PAD – peripheral artery disease, Afib – atrial fibrillation, GCC – glucocorticosteroid drugs, DMARD – disease-modifying antirheumatic drug.	
6 7	202		
8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41	202	To May 221 Downleaded from http://tomjopen.bmj.com/ on April 17, 2024 by guest. Protected by copyright	
42 43 44		For peer review only - http://bmjopen.bmj.com/site/about/guidelines.xhtml	11

BMJ Open

2 3 203	For three services, a major guideline relevant in 2014 in Switzerland could not be identified (long-term
4 5204	use of proton pump inhibitors, minor inpatient surgery procedures, elective Caesarean section). A total of eight
6 7205 8	services had weak, and six services had negative underlying recommendations. Median MOR was 1.29 for
9 206 10	services with weak and 1.25 for services with strong recommendations; 1.26 for services with positive and 1.46
11 207	for services with negative recommendations (Figure 1).
13 12408	
15 12609	Figure 1 Geographic variation of the health care services grouped by strength (A) and direction (B) of
17 1 2810 19	recommendations
²⁰ 11 21	A Weak and strong recommendations; B Positive and negative recommendations. MOR – median odds ratio.
22 2312	Boxplots depict the interquartile range of values (upper and lower hinges), and the median value.
24 2 2 13	
26 2 27 14	Based on Welch's t-test, the difference in mean variances [95Cl%] of services with weak and strong
28 29 215 30	recommendations was 0.03 [-0.06, 0.11], and the difference in mean MOR was 0.05 [-0.11, 0.21]. The difference
$31 \\ 32 \\ 32 \\ 32 \\ 32 \\ 32 \\ 32 \\ 32 \\ $	in mean variances [95CI%] of services with negative and positive recommendations was 0.07 [-0.03, 0.18], and
33 3 2 417	the difference in mean MOR was 0.14 [-0.06, 0.34].
35 32618	
37 38	7
38 3919 40	Discussion
41 42 42	We did not find a direct association between the strength of clinical recommendation and the geographic
43 4 2 421	variation in the utilization of 24 health care services. The geographic variation in the utilization of services with
45 4252 2	underlying negative recommendations was slightly higher than for those with positive recommendations, and
47 4223 49	for services with underlying weak recommendations than for those with strong recommendations. The
50 51 51 51	difference was larger for negative vs positive recommendations; however, both differences were not statistically
52 5 <u>3</u> 25	significant. In general, moderate potentially unwarranted geographic variation was observed, with MOR smaller
54 5 2 526	than 1.50 for all but one service.
56 5 2/27 58	At least two other studies have to some extent examined the association between the strength of
⁵⁹ 228	recommendations and the variation in adherence, each focusing on a single clinical specialty. In et al [7],
229	examining a set of recommendations in oncology, found higher variation in the utilization of services associated

with a lower level of evidence. However, this study focused not on regional but on inter-institutional variation,
 comparing two groups of providers. In contrast to this study and in agreement with our results, Mayer et al [8]
 found that surgeon practices in knee and hip arthroplasty in Australia varied regardless of the strength of
 evidence available. Potentially, different clinical areas could be associated with different barriers to guideline
 implementation, modifying the relationship between recommendations and variation.

To better understand why a direct association of recommendation strength and variation in adherence was not observed, it may be useful to revisit the EBM framework [1]. The EBM framework is *normative* and defines how clinical decisions *should* be made [1, 29]. However, this may not always coincide with how decisions *are* made – a process analysed by *descriptive* theories [29]. In fact, the EBM model has been developed as a conceptual rather than practical guidance of evidence implementation [1], and has not yet generated a coherent theory of clinical decision making, and in particular, of how evidence is incorporated [30]. Thus, although a direct relationship between the strength of recommendation and the geographic variation of service utilization would be encouraged by the normative EBM framework, it may not always be observed.

There are numerous reasons why even strong recommendations [31, 32], or conclusive research evidence more broadly [33], may not directly translate into clinical practice. Research on knowledge translation has identified multiple barriers at different levels of the health care system, including structural, organizational, peer-group, and professional factors [34] – many of which depend on the specific context where a service is provided, and thus may vary geographically. Knowledge transfer processes are highly non-linear and rely on triggering mechanisms [35]. Factors external to research evidence significantly affect translation – potentially creating large geographic heterogeneity even within services with strong recommendations. Finally, strong recommendations sometimes describe care with varying patient preferences. For example, although colon cancer screening is strongly recommended, patient preferences for test attributes and modalities vary significantly [36, 37].

An influential framework, explaining different degrees of variation between health care services, has been proposed by Wennberg and colleagues [38]. According to this framework, services are classified into effective, preference-sensitive, and supply-sensitive care. Effective care (services based on solid evidence, so that virtually all patients would choose them) largely corresponds to services with strong recommendations, as

defined by GRADE and applied in this study [2]. Preference-sensitive care partly corresponds to services with
weak recommendations, as they both imply trade-offs of risks and benefits of multiple options of care [38]. The
utilization of preference-sensitive surgical procedures usually has higher variation than of those associated with
effective care [39]. In contrast, supply-sensitive care defines the frequency, setting, and intensity of care
provision rather than specific types of health care services. It is associated with high, supply-related variation,
but is rarely discussed in guidelines [39], and therefore, could not be included in our study. However, the service
of minor surgeries performed as inpatient instead of outpatient procedures could be considered close to the
supply-sensitive category. In fact, it had the highest MOR (1.68) in our study.

Regarding the secondary hypothesis, we found that services associated with negative recommendations had slightly higher geographic variation. We did not find other studies directly comparing the regional variation of services with the direction of recommendations. Few studies, focusing mostly on low-value care, have reported MOR as an expression of geographic variation, further limiting the comparison. For example, in a study by Badgery-Parker et al [40], services discouraged by *Choosing Wisely* were shown to have regional MOR from 1.1 to 2.6 – a range that includes all of our observed MORs.

Negative recommendations usually address a widespread service that lacks supporting evidence of benefit or the benefit is outweighed by harms [2]. In contrast to services with positive recommendations, which are introduced after supporting evidence is produced, services with negative recommendations typically become part of the clinical practice *before* evidence is sufficient to rule out their overall benefit. Therefore, their use could be related more to clinical expertise and practice, and could be expected to vary locally. Indeed, the barriers to implementing positive and negative recommendations seem to be different [15] – signalling that the pathways how they are interpreted and integrated into clinical decisions might also be different. As Choosing Wisely and similar initiatives are increasingly gaining attention [41], our finding of higher geographic variation associated with negative recommendations may inform future research and implementation strategies.

This exploratory study has several limitations. First, although we aimed at a balanced selection of clinical fields and service types, the number (24) and range of studied services was limited by the data source, leading to somewhat unbalanced groups of strong and weak, positive and negative recommendations. Swiss claims data lack information on outpatient diagnoses, inpatient treatment details, and clinical information such as test

results [18]. Lack of clinical information also meant that some populations were not as specific as defined by the recommendation. For example, beta-blockers and angiotensin-converting-enzyme (ACE) inhibitors are recommended after a myocardial infarction contingent on heart failure and left ventricular dysfunction. As such clinical details were unavailable, we had to rely on them being present in the majority of the hospitalized myocardial infarction cases and distributed equally geographically. However, we believe that estimates of variation are accurate, as each of the 24 data points was generated by multilevel modelling of utilization in populations of 85000 patients on average, including all major explanatory variables such as age, sex, and indicators of morbidity. Second, the services studied were unavoidably different by characteristics other than the strength or direction of recommendation, such as service type or clinical area, potentially resulting in confounding. Indeed, although distributed among all recommendation types, diagnostic services had somewhat higher regional variation in utilization compared to treatment services (see Supplementary file 5). Although most of the selected services are delivered by primary care providers, their varied nature also means that the applied MobSpat regional units might not capture the regional variation equally well. Third, both the observed utilization and its geographic variation depend on the definition of the service and population [42]. We aimed to measure the *unwarranted* variation in utilization by using service-specific denominators (eligible populations) and adjusting for relevant clinical characteristics. How exactly *unwarranted* and *warranted* variation should be defined and measured, and what adjustments are necessary to differentiate them, is debated [22, 43]. Fourth, the grouping of recommendations by strength and direction was partly subjective, although we tried to make it reproducible with a clear algorithm, implemented in duplicate. Unfortunately, many different systems for evaluating the strength of recommendations exist [44], which cannot be easily reconciled, and the most prominent, GRADE approach, is not always explicitly used.

To explore the studied questions further, the sample of services could be expanded to inpatient and specialist care. Further, a meta-study of the numerous individual studies of geographic variation in health care services could be undertaken. However, this would currently be challenging, as studies choose different adjustment variables and specificity of studied populations, and report the variation in different quantitative forms (e.g., MOR, systematic component of variation, range). Furthermore, there is a need for qualitative studies of the reasons for the variability of clinical decisions, and how clinical expertise in these decisions

BMJ Open

interacts with evidence, clinical circumstances, and patient preferences. Qualitative evidence could help to
 generate more complex hypotheses for further quantitative studies, built on a finer understanding of all the
 factors influencing the variability of clinical decisions. Specialty-specific sets of services could also be further
 investigated.

L6 Conclusions

In this exploratory study of 24 health care services mostly in the outpatient primary care setting, we did not observe a significant difference in the degree of geographic variation in utilization of services associated with strong versus weak recommendations. Services associated with negative recommendations had slightly, although also not statistically significantly, higher geographic variation. The relationship between the strength of recommendations and the variation may be indirect or significantly modified by other characteristics of services, such as service type or clinical area. As initiatives discouraging low-value care are gaining attention worldwide, these findings may inform future research in this area.

5 Funding

This work was supported by the Swiss National Science Foundation (SNSF) National Research Program "Smarter
Health Care" (NRP 74), as part of project number 26, grant number 407440 167349.

29 **Competing interests**

MS declares a grant from Helsana Insurance Group, outside the submitted work. Helsana Group provided support in the form of salaries for authors BB, EB and CB, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The other authors declare no competing interests.

5 Authors' contributions

MS, VvW and HD developed the underlying study program. AU and HD developed study design, with support from all authors. AU, WW, CB, EB, BB did data extraction, preparation and management. WW, MS and OG 2 338 developed multilevel models applied in this study. AU and WW analysed the data. AU drafted the manuscript, **3**39 with major inputs from HD and MS, and contributions from all authors. All authors read and approved the final 7340 manuscript. 8 9341 10 Data sharing statement The data underlying this study cannot be shared publicly because they are the property of Helsana 15 13644 (https://www.helsana.ch/en/helsana-group), and have restricted public access on grounds of patient privacy. 17 13845 The data are managed by Helsana and subsets of the database are available for researchers after request and 19 under specific conditions. Data are available from Helsana (gesundheitskompetenz@helsana.ch) for researchers who meet the criteria for access to confidential data. Helsana will consider the possibilities of the research proposal and decide to grant access if the research questions can be answered with use of the Helsana data. 23:48 26 23749 When requests are granted, data are accessible only in a secure environment. 28 31 32251 References 33 34 3352 1. Haynes RB, Devereaux PJ, Guyatt GH. Clinical expertise in the era of evidence-based medicine and patient 36 3753 choice. Evid Based Med. 2002;7:36-8. 38 2. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from evidence to recommendations: The significance and presentation of recommendations. J Clin Epidemiol.

4356 2013;66:719–25. doi:10.1016/j.jclinepi.2012.03.013.

43

45

52

457 3. Andrews JC, Schünemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going

58 from evidence to recommendation - Determinants of a recommendation's direction and strength. J Clin

59 Epidemiol. 2013;66:726–35. doi:10.1016/j.jclinepi.2013.02.003.

4. Cutler D, Skinner JS, Stern AD, Wennberg D. Physician beliefs and patient preferences: A new look at regional

3561 variation in health care spending. Am Econ J Econ Policy. 2019;11:192–221. 56

5. Finkelstein A, Gentzkow M, Williams H. Sources of Geographic Variation in Health Care: Evidence From

363 PatientMigration. Q J Econ. 2016;131:1681–726.

364 6. Molitor D. The evolution of physician practice styles: Evidence from cardiologist migration. Am Econ J Econ

2 365 3	Policy. 2018;10:326–56.
4 5366	7. In H, Neville BA, Lipsitz SR, Corso KA, Weeks JC, Greenberg CC. The role of National Cancer Institute-
o 7367 8	designated cancer center status: observed variation in surgical care depends on the level of evidence. Ann Surg.
9 368 10	2012;255:890–5. doi:10.1097/SLA.0b013e31824deae6.
11 369 12	8. Mayer M, Naylor J, Harris I, Badge H, Adie S, Mills K, et al. Evidence base and practice variation in acute care
13 1 <mark>3</mark> 70	processes for knee and hip arthroplasty surgeries. PLoS One. 2017;12:e0180090.
15 1 3671 17	doi:10.1371/journal.pone.0180090.
1 3872 19	9. Corallo AN, Croxford R, Goodman DC, Bryan EL, Srivastava D, Stukel TA. A systematic review of medical
²⁰ 373 21	practice variation in OECD countries. Health Policy (New York). 2014;114:5–14.
22 374 23	doi:10.1016/j.healthpol.2013.08.002.
24 2 3 75 26	10. The Dartmouth Atlas of Health Care. https://www.dartmouthatlas.org/. Accessed 11 May 2020.
20 2 3 76 28	11. Public Health England. Atlas of Variation. https://fingertips.phe.org.uk/profile/atlas-of-variation. Accessed
²⁰ 377 30	11 May 2020.
31 378 3278	12. Cassel CK, Guest JA. Choosing wisely: Helping physicians and patients make smart decisions about their care.
33 3 3 479 35	JAMA - Journal of the American Medical Association. 2012;307:1801–2.
33580 37	13. Elshaug AG, Watt AM, Mundy L, Willis CD. Over 150 potentially low-value health care practices: an Australian
³ 881 39	study. Med J Aust. 2012;197:556–60.
40 382	14. Prasad V, Vandross A, Toomey C, Cheung M, Rho J, Quinn S, et al. A decade of reversal: An analysis of 146
42 43383	contradicted medical practices. Mayo Clin Proc. 2013;88:790–8.
44 4 5 84 46	15. Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: A meta-synthesis of GPs' attitudes to clinical
4 3 85 48	practice guidelines. Br J Gen Pract. 2007;57:971–8.
49 386 50	16. National Research Programme 74. Project 26: How do guidelines and recommendations influence medical
51 5 3 87	treatment? http://www.nfp74.ch/en/projects/healthcare-across-sectors/project-schwenkglenks. Accessed 15
53 5 3488 55	Apr 2020.
5 5 89 57	17. Swiss Federal Office of Public Health and Swiss Conference of Cantonal Health Directors. Nationale Strategie
5890 5990	Prävention nichtübertragbarer Krankheiten (NCD-Strategie) 2017-2024 [National Strategy of Non-communicable
60 391	Diseases]. 2016.

¢

1

2 392	18. Ulyte A, Bähler C, Schwenkglenks M, von Wyl V, Gruebner O, Wei W, et al. Measuring diabetes guideline
4 5393	adherence with claims data: systematic construction of indicators and related challenges. BMJ Open.
6 7394 8	2019;9:e027138. doi:10.1136/bmjopen-2018-027138.
9 395 10	19. Bachmann L, Ulyte A, Dressel H. Clinical practice guidelines of medical societies in Switzerland: analysis of
11 396 12	the current state. Swiss Med Wkly. 2019;149:w20134.
13 1 3 97 15	20. Guyatt GH, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. GRADE: Going from Evidence to
13698 17	Recommendations. BMJ. 2008;336:1048–51.
1 3 899 19	21. Bundesamt für Statistik. MS-Regionen. https://www.bfs.admin.ch/bfs/de/home/statistiken/raum-
2000 21	umwelt/nomenklaturen/msreg.assetdetail.415729.html. Accessed 11 May 2020.
22 401 24	22. Sutherland K, Levesque J. Unwarranted clinical variation in health care: Definitions and proposal of an
24 24502 26	analytic framework. J Eval Clin Pract. 2019;:jep.13181. doi:10.1111/jep.13181.
2 4703 28	23. Larsen K, Petersen JH, Budtz-Jørgensen E, Endahl L. Interpreting parameters in the logistic regression model
²² 04 30	with random effects. Biometrics. 2000;56:909–14.
31 405 32 33	24. Larsen K, Merlo J. Appropriate Assessment of Neighborhood Effects on Individual Health: Integrating
3406 35	Random and Fixed Effects in Multilevel Logistic Regression. Am J Epidemiol. 2005;161:81–8.
3 4607 37	25. Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-
³⁴ 08 39	Whitney U test. Behav Ecol. 2006;17:688–90.
40 4109 41	26. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical
42 ∡∯10 44	Computing, Vienna, Austria. 2020. https://www.r-project.org/. Accessed 17 Dec 2020.
4511 46	27. Charlton, C, Rasbash, J, Browne, WJ, Healy, M, Cameron B. MLwiN Version 3.01. Centre for Multilevel
4712 48	Modelling, University of Bristol. 2017.
49 413	28. Leckie G, Charlton C. Runmlwin: A program to run the MLwiN multilevel modeling software from within
51 <u>54</u> 14 53	Stata. J Stat Softw. 2013;52:1–40.
54415 55	29. Djulbegovic B, Elqayam S, Dale W. Rational decision making in medicine: Implications for overuse and
5 4 916 57	underuse. J Eval Clin Pract. 2018;24:655–65. doi:10.1111/jep.12851.
58 59 59	30. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. The Lancet.
60 418	2017;390:415–23.

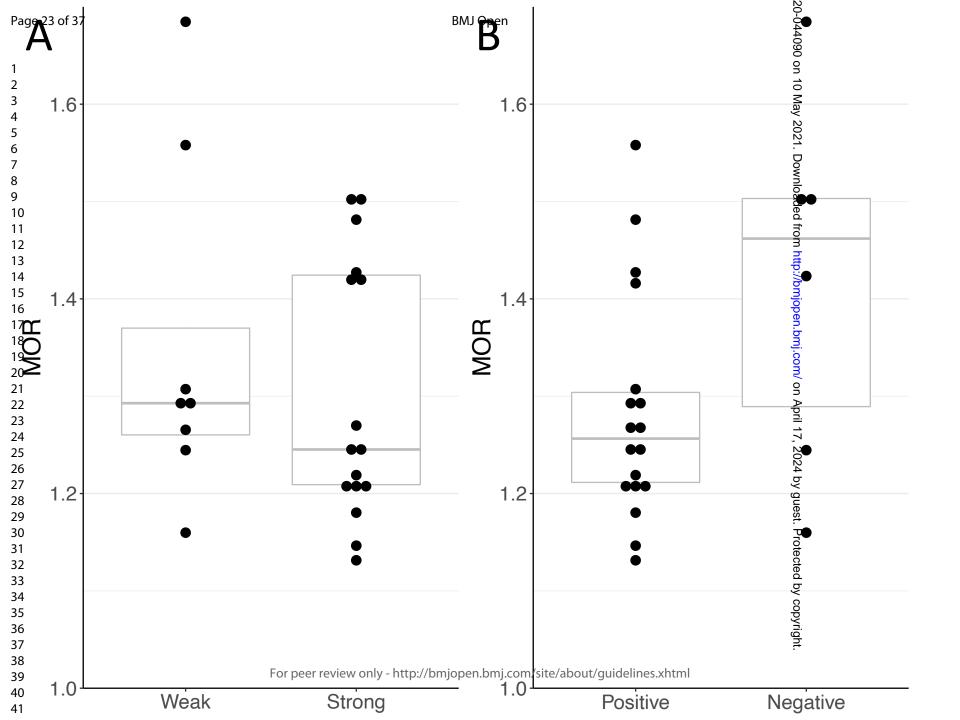
1	
2419 3	31. Francke AL, Smit MC, De Veer AJE, Mistiaen P. Factors influencing the implementation of clinical guidelines
4 ₅420	for health care professionals: A systematic meta-review. BMC Med Inform Decis Mak. 2008;8:38.
6 7421 8	doi:10.1186/1472-6947-8-38.
9 422 10	32. Lugtenberg M, Zegers-Van Schaick JM, Westert GP, Burgers JS. Why don't physicians adhere to guideline
11 423 12	recommendations in practice? An analysis of barriers among Dutch general practitioners. Implement Sci.
13 1424	2009;4:1–9.
15 1 4625 17	33. Wallace J, Nwosu B, Clarke M. Barriers to the uptake of evidence from systematic reviews and meta-
14 26 19	analyses: A systematic review of decision makers' perceptions. BMJ Open. 2012;2:e001220.
20 24 27 21	34. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement
22 4328	Sci. 2012;7:50. doi:10.1186/1748-5908-7-50.
24 2 4529 26	35. Braithwaite J, Churruca K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation
20 2 4730 28	science: A theoretical and empirical analysis of systems change. BMC Med. 2018;16:63. doi:10.1186/s12916-
² 431 30	018-1057-z.
31 432 32	36. Marshall DA, Johnson FR, Phillips KA, Marshall JK, Thabane L, Kulin NA. Measuring Patient Preferences for
33 3 4 33 35	Colorectal Cancer Screening Using a Choice-Format Survey. Value Heal. 2007;10:415–30.
35 34634 37	37. Mansfield C, Tangka FKL, Ekwueme DU, Smith JL, Guy GP, Li C, et al. Stated Preference for Cancer Screening:
3435 39	A Systematic Review of the Literature, 1990-2013. Prev Chronic Dis. 2016;13:E27. doi:10.5888/pcd13.150433.
40 4136	38. Wennberg JE, Fisher ES, Skinner JS. Geography and the debate over medicare reform. Health Affairs. 2003;22
42 4 3 37	SUPPL.
44 4 5 38 46	39. Wennberg JE. Tracking medicine : a researcher's quest to understand health care. Oxford University Press;
4739 48	2010.
49 440 50	40. Badgery-Parker T, Feng Y, Pearson S-AA, Levesque J-FF, Dunn S, Elshaug AG. Exploring variation in low-value
51 5441	care: a multilevel modelling study. BMC Health Serv Res. 2019;19:345. doi:10.1186/s12913-019-4159-1.
53 54142 55	41. Levinson W, Kallewaard M, Bhatia RS, Wolfson D, Shortt S, Kerr E a. "Choosing Wisely": a growing
55 54643 57	international campaign. BMJ Qual Saf. 2015;24:167–74. doi:10.1136/bmjqs-2014-003821.
58 444 59	42. Schwartz AL, Landon BE, Elshaug AG, Chernew ME, McWilliams JM. Measuring low-value care in Medicare.
⁶⁰ 445	JAMA Intern Med. 2014;174:1067–76. doi:10.1001/jamainternmed.2014.1541.

¢

43. Harrison R, Manias E, Mears S, Heslop D, Hinchcliff R, Hay L. Addressing unwarranted clinical variation: A
rapid review of current evidence. J Eval Clin Pract. 2019;25:53–65. doi:10.1111/jep.12930.
44. Atkins D, Eccles M, Flottorp S, Guyatt GH, Henry D, Hill S, et al. Systems for grading the quality of evidence
and the strength of recommendations I: Critical appraisal of existing approaches. BMC Health Serv Res.
2004;4:1–7.

Supplementary files

Supplementary file 1 Definitions and descriptions of the studied health care services and eligible populations Supplementary file 2 Algorithm and criteria for the assessment of the strength of recommendation Supplementary file 3 Algorithm and criteria for the assessment of the direction of recommendation Supplementary file 4 List of guidelines selected for the study, describing the services analysed Supplementary file 5 Regional variation of diagnostic and treatment services studied



Additional file 1 Definitions and descri	iptions of the studied health care	e services and eligible populations

				Recommendation for the	Specific clinical explanator	Clinical codes used for identification
	service	and frequency	population	health care service	variables	of the health care service
creening	Colon cancer	Colonoscopy/ year	Anyone 50-69	Colonoscopy should be done	Previous treatment of cancer or	Colonoscopy: 19.06 (TM Kapitel);
	screening		years old	every 10 years for people 50-69	inflammatory bowel diseas	G48% (DRG);
				years old.	hospitalization with colon deease	45.23, 45.25, 48.29.1%, 48.29.2%
			\sim		in the last year	(CHOP)
					ownl	
			0r		Downloaded	
					d fro	
-	Breast cancer	Mammography/ year	50-74 years old	Mammography should be done	Previous treatment of brease or	Mamography: 39.1310, 39.1320,
	screening		women	every 2 years for 50-74 years	other cancer tp://bmjop	39.1307, 39.1308, 39.1300, 39.1305
				old women.	omjo	39.1306 (TM);
				· .	ben.t	TZ
	Prostate cancer	Prostate-specific	50-70 years old	Early detection of prostate	Previous treatment of cancer,	PSA testing: 1626.00 (Ana)
	screening	antigen (PSA)	men	cancer (opportunistic	hospitalization with prostate	
		testing/ year		screening) should be offered to	disease in the last year $\stackrel{G}{\triangleright}$	
				the well-informed man.	pri 1	
	Osteoporosis	Dual-energy x-ray	Women over 60	DXA densitometry is	Presence of more than one risk	DXA densitometry: 39.1950, 39.214
	screening	absorptiometry	and with risk	recommended for	factor 24	39.2150, 39.2160 (TM)
		(DXA)/ year	factors ^a of	postmenopausal women with	n ĝ v	
			spontaneous	spontaneous fractures or	est. Prote	
	·	absorptiometry	and with risk factors ^a of	recommended for postmenopausal women with	factor 024 by guest.	

BMJ Open

Pag	e 25 of 37				BMJ Open	6/bmjopen	
1 2 3 4 5 6	Diagnosis	DM: HbA1c test	Glycated haemoglobin (HbA1c) test twice/ year	>18-year-old drug-treated diabetes patients	HbA1c test should be done for diabetes patients at least twice a year.	Oral diabetes medication of insulin	HbA1c test: 1363.00, 1363.01 (Ana)
7 8 9 10 11 12 13		DM: renal function test	Albuminuria and serum creatinine tests/ year	>18-year-old drug-treated diabetes patients	Albuminuria and serum creatinine tests should be done for diabetes patients at least once a year.	Oral diabetes medication or insulin	Albuminuria: 1023.00, 1023.01, 1739.00, 1739.01, 1740.00, 1740.01 (Ana) Serum creatinine: 1509.00, 1509.01 (Ana)
14 15 16 17 18 19 20 21 22 23		DM: LDL test	Low-density lipoprotein (LDL) test/ year	19-75-year-old drug-treated diabetes patients	LDL test should be done for diabetes patients at least once a year.	Oral diabetes medication of insulin	LDL test: 1521.00 (Ana) Total cholesterol test: 1230.00, 1230.01 (Ana) HDL test: 1410.01, 1410.10 (Ana) Triglycerides test: 1731.01, 1731.00 (Ana)
24 25 26 27 28 29		DM: eye examination	Ophthalmologist visit/ year	>18-year-old drug-treated diabetes patients	Eye exam should be performed for diabetes patients at least once a year.	Oral diabetes medication of insulin Ap	Outpatient visit with ophthalmologist: (sub group "Ophthalmologie" in Swiss care provider registry sasis.ch)
30 31 32 33 34 35 36 37		TSH screening	Thyroid-stimulating hormone (TSH) test without T3 and T4 tests on the same day	>18-year-old persons without thyroid disease ^b and receiving TSH test	TSH should be measured as an initial screening test for hypo/hyperthyroidism, while T3 and T4 test should follow if TSH is abnormal.	-	TSH test: 1718.10 (Ana) T3 or T4 test: 1732.00, 1720.00, 733.00, 1721.00 (Ana)
38 39 40 41 42						d by copyright.	

OCR nfluenza accination	Outpatient preoperative chest radiography (POCR) up to 2 months before surgery Influenza outpatient vaccination/ year	>18-year-old patients with inpatient surgical procedures People over 65 years old or with a specified	Routine chest radiography is not recommended before surgery. People over 65 years old and patients with chronic conditions, specified by Federal	2020-044090 on 10 Mayonia Hospitalization with pneumonia in the last year	Chest radiography: 39.0190 (TM) Influenza vaccination: J07BB02 (ATC)
	radiography (POCR) up to 2 months before surgery Influenza outpatient	inpatient surgical procedures People over 65 years old or with	surgery. People over 65 years old and patients with chronic	Hospitalization with pneumonia	Influenza vaccination: J07BB02 (ATC)
	up to 2 months before surgery Influenza outpatient	surgical procedures People over 65 years old or with	People over 65 years old and patients with chronic	Hospitalization with pneumonia	Influenza vaccination: J07BB02 (ATC)
	before surgery Influenza outpatient	procedures People over 65 years old or with	patients with chronic	Hospitalization with pneumonia	Influenza vaccination: J07BB02 (ATC)
	Influenza outpatient	People over 65 years old or with	patients with chronic	Hospitalization with pneumonia	Influenza vaccination: J07BB02 (ATC)
	-	years old or with	patients with chronic	in the last year	Influenza vaccination: J07BB02 (ATC)
accination	vaccination/ year				
		a specified	conditions specified by Federal		
			conditions, specified by rederal	Jow Sow	
		chronic	Office of Public Health, should	Downloaded	
		condition ^c	be vaccinated against influenza		
		\mathcal{O}	every year.	from	
enzodiazepines	Cumulative	Anyone over 65	Long-term use of	Treated epilepsy, stay in a	Benzodiazepines and other hypnotics:
	prescription of	years old	benzodiazepines and other	nursing home in the last ye $\overline{\underline{a}}$,	N03AE01, N05BA%, N05CD%, N05BB%,
	benzodiazepines		hypnotics is discouraged for old	hospitalization in the last year	N05BE%, N05CA%, N05CB%, N05CC%,
	(BZD) for >8 weeks/		patients.	with a diagnosis indicative of	N05CF%, N05CH%, N05CM%, N05CX%
	year		6	justified benzodiazepine use	(ATC)
roton pump	Cumulative	>18-year-old	PPI should not be used at	- <u>O</u>	PPI or H2: A02BC%, A02BD%,
nhibitors	prescription of	persons	maximal dose for prolonged	April	M01AE52, A02BA% (ATC)
	proton pump	receiving PPI or	periods of time.	17, 2	
	inhibitors (PPI) or H2	H2 drugs		2024	
	histamine receptor			by g	
	antagonists (H2) for			Juest	
	>8 weeks/ year			. Pro	
npatient	Specified surgical	>18-year-old	If none of the special conditions	- te	
rocedures	procedures ^d done in	patients with	apply, certain surgical	d by	
np	oton pump ibitors patient	prescription of benzodiazepines (BZD) for >8 weeks/ year toton pump Cumulative prescription of proton pump inhibitors (PPI) or H2 histamine receptor antagonists (H2) for >8 weeks/ year	prescription of years old benzodiazepines (BZD) for >8 weeks/ year boton pump Cumulative >18-year-old prescription of persons proton pump receiving PPI or inhibitors (PPI) or H2 H2 drugs histamine receptor antagonists (H2) for >8 weeks/ year batient Specified surgical >18-year-old	prescription of benzodiazepines (BZD) for >8 weeks/ yearyears oldbenzodiazepines and other hypnotics is discouraged for old patients.oton pumpCumulative prescription of proton pump inhibitors (PPI) or H2 histamine receptor antagonists (H2) for >8 weeks/ year>18-year-old persons receiving PPI or H2 drugsPPI should not be used at maximal dose for prolonged periods of time.histamine receptor antagonists (H2) for >8 weeks/ yearH2 drugsH2 drugsbatientSpecified surgical>18-year-oldIf none of the special conditions	prescription of benzodiazepines (BZD) for >8 weeks/ yearyears oldbenzodiazepines and other hypnotics is discouraged for old patients.nursing home in the last year hospitalization in the last year with a diagnosis indicative of justified benzodiazepine useoton pump libitorsCumulative prescription of proton pump inhibitors (PPI) or H2 histamine receptor antagonists (H2) for >8 weeks/ year>18-year-old periods of time.PPI should not be used at maximal dose for prolonged periods of timepratientSpecified surgical>18-year-oldIf none of the special conditions-

Page	e 27 of 37				BMJ Open	6/bmjopen-	
1			the outpatient	specified	procedures should be done in		
2 3			setting	surgical	the outpatient setting.	0-04	
4				procedures		4090	
5 6				(either as in- or		0 on	
7				outpatient)		2020-044090 on 10 May 2021. Downloaded from	
8 9		Caesarean	Caesarean section	>18-year-old	C-section should not be	- 20	C-section: 74.0%, 74.1%, 74.2%, 74.4%,
10		section	(C-section)	women giving	performed unless absolute or	021.	74.99 (CHOP); O01A, O01B, O01C,
11 12				birth without	relative indications are present.	Dow	001D, 001E, 001F (DRG); 22.2120,
13				absolute		nloac	22.2130, 22.2410, 22.2420 (TM)
14 15				indications ^e for		ded f	
16 17				C-section	0	rom F	
18	Secondary	AMI: aspirin	Aspirin prescription	>18-year-old	All myocardial infarction	Hospitalization for stroke of	Aspirin: B01AC06 (ATC)
19 20	prevention		within 2 weeks after	patients with	patients should take aspirin	bleeding event or prescribe	
21			acute myocardial	AMI ^f	long-term.	anticoagulation in the last year	
22 23			infarction (AMI)		Via	.bmj.	
24		AMI: statin	High-dose statin	>18-year-old	All myocardial infarction	Hospitalization for stroke in the	High-dose statins: C10AA05, C10AA07
25 26			prescription within 2	patients with	patients should get statins long-	last year 9	(ATC)
27 28			weeks after AMI	AMI ^f	term.	April	
28 29		AMI: beta-	Beta-blocker	>18-year-old	All myocardial infarction	Hospitalization with heart failure	Beta-blockers: C07% (ATC)
30 31		blocker	prescription within 2	patients with	patients with heart failure or	diagnosis in the last year $\overset{10}{24}$	
32			weeks after AMI	AMI ^f	impaired function should get	by g	
33 34					beta-blockers long-term.	guest.	
35		AMI: ACE/ARB	Angiotensin-	>18-year-old	All myocardial infarction		ACE or ARB medication: C09% (ATC)
36 37			converting enzyme	patients with	patients with heart failure or	Protected by	
38			(ACE) or angiotensin	AMI ^f	impaired function should get		
39 40		1	1	1	1	dopyright.	
41						right.	
42							

Page	28	of	37
------	----	----	----

	receptor blocker		ACE or ARB antihypertensive	.202	
	(ARB)		medication long-term.	0-04	
	antihypertensive			409	
	medication			0 on	
	prescription within 2			10 N	
	weeks after AMI			1ay 2	
AMI: P2Y12	P2Y12 antiplatelet	>18-year-old	All myocardial infarction	6/bmjopen-2020-044090 on 10 May 202 Hospitalization for a bleeding	P2Y12 drugs: B01AC04, B01AC22
inhibitors	drug ^g prescription	patients with	patients should get P2Y12	event or prescribed	B01AC24 (ATC)
	within 2 weeks after	AMI	antiplatelet drugs for at least 1-	anticoagulation in the last verain	
	AMI		12 months according to the	ded t	
			bleeding risk profile and AMI	ded from http	
			treatment.	http:/	
PPI with NSAID	PPI prescription	>18-year-old	Patients taking long-term NSAID	Concurrent use of antiplate	NSAID: M01A% (ATC)
	within 1 month or up	patients with a	and with risk factors for gastric	anticoagulation drugs or or	PPI: A02BC%, A02BD%, M01AE5
	to 3 months before	cumulative	ulcer ^h should also take PPI.	glucocorticoids, hospitalization	(ATC)
	initial long-term	NSAID	0	for bleeding event in the las	
	nonsteroidal anti-	prescription of		year. 9	
	inflammatory drug	>8 weeks at		April	
	(NSAID) prescription	maximal dose		on April 17, 2024 by gues	
PAD: statin	Prescription of	>18-year-old	Statins are recommended for all		Statins: C10AA%, C10B% (ATC)
	statins within 3	patients	patients with PAD.	by g	
	months after	undergoing			
	peripheral artery	diagnostic or		t. Protected by copyright.	
	disease (PAD)	treatment		tecte	
	identification			d b	

9 of 37				BMJ Open		6/bmjopen-	
			procedures for PAD ⁱ			- <u>2</u> 020-044090 on 10 May 2021.	
Δf	fib:	Oral anticoagulation	>18-year-old	All patients with atrial		4409	Oral anticoagulation: B01AE07,
	nticoagulation	prescription within 2	patients with	fibrillation should be prescribed		õ on	B01AF01, B01AF02, B01AF03,
u	licougulation	weeks after atrial	atrial fibrillation	oral anticoagulation for embolic		101	B01AA04, B01AA07 (ATC)
		fibrillation (Afib)	diagnosis and	events prevention according to		May	DUIANO, DUIANO (AIC)
		identification	-			202	
		Identification	additional risk	the CHA ₂ DS ₂ -VASc score.		Г. Do	
			factors ⁱ			Downk	
	CC with new	Glucocorticoid (GCC)	>18-year-old	Short-term glucocorticoids	-	loaded from http://bmjopen.bmj.com/	Glucocorticoids: H02% (ATC)
DI	MARD	prescription within 1	patients with a	should be taken with newly		d fro	DMARD: L01BA01, L04AX03,
		month or up to 3	new prescription	prescribed DMARD.		m h	M01CX01, L04AA13, M01CX02,
		months before	of DMARD by a	~ 0×		ttp://	P1BA02, P01BA01, M01CC01,
		disease-modifying	rheumatologist	the second se		bmjc	L01AA01, M01CB01, L04AX01 (ATC)
		antirheumatic drug		· · · · ·		pen.	
		(DMARD)				bmj.	
		prescription		2	1.	com	
spondylitis malnutritio b. Hyperth	s, osteogenesis in on, hypogonadis nyroidism, hypot	mperfecta, rheumatoid a m, hyper- or hypothyroi hyroidism, goitre or thyr	arthritis, inflammato dism, and hyperpara roiditis.	e, use of drugs increasing the risk o ory bowel disease, Cushing's disease athyroidism. Patients currently trea c liver disease, renal failure, immur	e, alcohol or nicotine abuse, ited or diagnosed with ostec	rii Tronic 2024 by c	liver disease, gastrectomy, were excluded.
				rhoids, inguinal hernia and cervix, k		St -	
						0	, tonsmectomy.
			-	tion, or multiple pregnancy.		ecte	
-		n a diagnosis of acute my	ocardial infarction (AIVII).		d by	
g. Clopidoį	grel, prasugrel o	r ticagrelor.				tected by copyright.	
			For peer review o	nly - http://bmjopen.bmj.com/site			6

 i. Peripheral artery disease (PAD) or carotid stenosis diagnosed during an inpatient stay, amputation of lower or upper extremity, throgebectomy, stenting or other procedures in peripheral arteries, specialized diagnostic ultrasound, magnetic resonance tomography (MRI) angiography, computer tomography (CToongiography or angiography of peripheral arteries.

j. Risk factors (congestive heart failure, hypertension, age 65-74 or ≥75 years old, diabetes, previous stroke, transient ischemic attack, or thromboembolism, cardiovascular disease, female sex) were extracted from available claims data and summed according to CHA2DS2-VASc score. Patients with CHA2DS2. VASc score of ≥2 for males and ≥3 for females were included.

DM – diabetes mellitus, HbA1c – Glycated haemoglobin, LDL – low density lipid, TSH – thyroid-stimulating hormone, T3 and T4 – triiodethyronine and thyroxine, POCR – preoperative chest radiography, BZD – benzodiazepines, PPI – proton pump inhibitors, H2 – H2 histamine receptor antagonists, C-section – Caesarean section, AMI – acute myocardial infarction, ACE/ARB – angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers, NSAID – nonsteroidal ati-inflammatory drugs, PAD – peripheral artery disease, Afib – atrial fibrillation, GCC – glucocorticosteroid drugs, DMARD – disease-modifying antirheumatic drug. Ana – Analysenliste, Swiss outpatient laboratory test codes; ATC - Anatomical Therapeutic Chemical Classification System, code and quantity of a prescription drug; CHOP -Schweizerische Operationsklassifikation, a classification of inpatient procedures; DRG - Swiss Diagnosis Related Groups, a classification of inpatient cases, based on diagnoses,

procedures and other clinical information; ICD - International Classification of Diseases, 10th revision, German Modification, codes for Brimary and secondary diagnoses for each

hospitalization episode of an inpatient; TM – Tarmed, Swiss classification of outpatient procedures and services; TM Kapitel – Tarmed Shapter codes; TZ – Tarifziffer, further

codes representing reimbursement of screening services within cantonal breast cancer screening programs.

v on April 17, 2024 by guest. Protected by copyright

6/bmjopen-202

N of authors	Steps	
Single	1.	Identify the relevant medical societies for each selected health care service.
	2.	Look up the European or international medical societies' websites and journals, and
		identify relevant guidelines published before 2014. In addition, look up if Swiss federal
		legislation guidelines exist by 2014.
	3.	If none found, look up American medical society and identify relevant guidelines
		published before 2014.
	4.	If none found, consider the recommendation weak.
In duplicate	5.	Once the guideline and the recommendation statement are located, classify the
		recommendation into strong or weak ^a .
	-	Strong recommendation implies that the desirable effects of adherence to a
		recommendation outweigh the undesirable effects.
	-	That means that most informed patients would choose the recommended
		management and that clinicians can structure their interactions with patients
		accordingly.
	-	For clinicians, that would mean that most patients should receive the recommended
		course of action.
		For patients, that would mean that most people in such a situation would want the
		recommended course of action and only a small proportion would not; patients shoul
		request discussion if the intervention is not offered.
	-	Weak recommendation implies that the desirable effects of adherence to a
		recommendation probably outweigh the undesirable effects, but the guideline panel i
		less confident.
	-	Thus, a weak recommendation is conditional or optional, and means that patients'
		choices will vary according to their values and preferences, and clinicians must ensure
		that patients' care is in keeping with their values and preferences.
	-	For clinicians, that would mean that they should recognize that different choices will b
		appropriate for different patients and that they must help each patient to arrive at a
		management decision consistent with her or his values and preferences.
		For patients, that would mean that most people in such situation would want the
		recommended course of action, but many would not.
adapted from:	Guyatt G	H, Oxman AD, Kunz R, Falck-Ytter Y, Vist GE, Liberati A, et al. GRADE: Going from Evider
		/J 2008;336:1048–51.

N of authors	Steps	
In duplicate	1.	Once the guideline and the recommendation statement are located (see Additional file
		2), classify the recommendation into positive and negative.
	-	Positive recommendation encourages the use of a health care service in a given
		population.
	-	Negative recommendation discourages the use of a health care service in a given
		population (e.g., contains negative indicatory words, such as not, no, never)

to beet terien only

BMJ Open

Recommendation	Reference	Comment	
Colon cancer screening	[1]		
Breast cancer screening	[2]		
Prostate cancer screening	[3]		
Osteoporosis screening	[4]		
DM: HbA1c test	[5]		
DM: renal function test	[5]		
DM: LDL test	[5]		
DM: eye examination	[5]		
TSH screening	[6,7]		
POCR	[8]		
Influenza vaccination	[9]		
Benzodiazepines	[10]		
Proton pump inhibitors		Swiss national guideline since 2016 [11]	
Inpatient procedures	- 0	Swiss federal regulation exists from 2019 [12]	
Caesarean section	- ``	Swiss national guideline since 2015 [13]	
AMI: aspirin	[14,15]		
AMI: statin	[14,15]		
AMI: beta-blocker	[14,15]		
AMI: ACE/ARB	[14,15]		
AMI: P2Y12 inhibitors	[14,15]		
PPI with NSAID	[16]		
PAD: statin	[17]		
Afib: anticoagulation	[18]		
GCC with new DMARD	[19]		

Additional file 4 List of guidelines selected for the study, describing the services analysed

DM – diabetes mellitus, HbA1c – Glycated haemoglobin, LDL – low density lipid, TSH – thyroid-stimulating hormone, POCR – preoperative chest radiography, AMI – acute myocardial infarction, ACE/ARB – angiotensinconverting enzyme inhibitors or angiotensin II receptor blockers, PPI – proton pump inhibitors, NSAID – nonsteroidal anti-inflammatory drugs, PAD – peripheral artery disease, Afib – atrial fibrillation, GCC – glucocorticosteroid drugs, DMARD – disease-modifying antirheumatic drug.

- Lansdorp-Vogelaar I, Karsa L. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition – Introduction. Endoscopy. 2012 Sep 25;44(S 03):SE15–30.
- Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med. 2009 Nov 17;151(10):716.
- Heidenreich A, Bellmunt J, Bolla M, Joniau S, Mason M, Matveev V, et al. EAU Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Treatment of Clinically Localised Disease. Eur Urol. 2011 Jan;59(1):61–71.

2 3

4

5 6

7

8 9

10

11 12

13

14 15

16

17 18

19

20 21

22 23

24

25

26 27

28 29

30

31 32

33

34 35

36

37 38

39

40 41

42

43 44

45

46 47

48

49 50

51

52 53

54

55 56

57

58 59

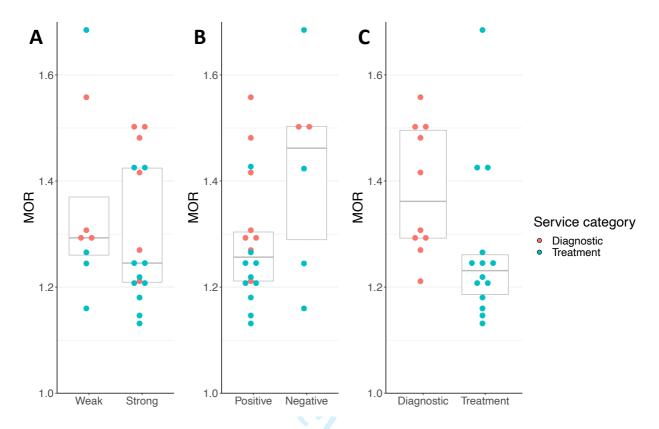
- 4. Kanis JA, McCloskey E V., Johansson H, Cooper C, Rizzoli R, Reginster J-Y. European guidance for the diagnosis and management of osteoporosis in postmenopausal women. Osteoporos Int. 2013 Jan 19;24(1):23-57. 5. Rydén L, Grant PJ, Anker SD, Berne C, Cosentino F, Danchin N, et al. ESC Guidelines on diabetes, prediabetes, and cardiovascular diseases developed in collaboration with the EASD. Eur Heart J. 2013 Oct 14;34(39):3035-87. 6. Bahn RS, Burch HB, Cooper DS, Garber JR, Greenlee MC, Klein I, et al. Hyperthyroidism and Other Causes of Thyrotoxicosis: Management Guidelines of the American Thyroid Association and American Association of Clinical Endocrinologists. Thyroid. 2011 Jun;21(6):593-646. Garber J, Cobin R, Gharib H, Hennessey J, Klein I, Mechanick J, et al. Clinical Practice Guidelines for 7. Hypothyroidism in Adults: Cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. Endocr Pract. 2012 Nov;18(6):988–1028. 8. Choosing Wisely. American College of Surgeons. Admission pre-op chest x-rays. September 4, 2013. Available from: https://www.choosingwisely.org/clinician-lists/american-college-surgeons-admission-orpreop-chest-x-ray-on-ambulatory-patients/ Accessed on May 11, 2020 9. Swiss Federal Office of Public Health. Recommendations for Influenza Vaccination (Empfehlungen zur Grippeimpfung). 2011. 10. Choosing Wisely. American Geriatrics Society. Benzodiazepines sedative hypnotics for insomnia in older adults. February 21, 2013. Available from: https://www.choosingwisely.org/clinician-lists/americangeriatrics-society-benzodiazepines-sedative-hypnotics-for-insomnia-in-older-adults/ Accessed on May 11, 2020 11. Swiss Society of General Internal Medicine. Smarter Medicine. Top-5 List for Ambulatory Care. 2016. Available from: https://www.smartermedicine.ch/de/top-5-listen/ambulante-allgemeine-inneremedizin.html Accessed on May 11, 2020 12. Swiss Federal Office of Public Health. "Outpatient instead of Inpatient" [Änderung der Krankenpflege-Leistungsverordnung (KLV) betreffend «Ambulant vor Stationär»]. Available from: https://www.bag.admin.ch/bag/de/home/versicherungen/krankenversicherung/krankenversicherungrevisionsprojekte/konsultation-ambulant-vor-stationaer.html Accessed on May 11, 2020 Hoesli I, Alma-Stucki S El, Drack G, Girard T, Irion O, Schulzke S, et al. Guideline Sectio Caesarea. 2015;1–20. 13. Available from: https://www.sggg.ch/fachthemen/guidelines/ Accessed on May 11, 2020 14. Hamm CW, Bassand J-P, Agewall S, Bax J, Boersma E, Bueno H, et al. ESC Guidelines for the management of acute coronary syndromes in patients presenting without persistent ST-segment elevation: The Task Force for the management of acute coronary syndromes (ACS) in patients presenting without persistent STsegment elevatio. Eur Heart J. 2011 Dec 1;32(23):2999-3054. 15. Steg PG, James SK, Atar D, Badano LP, Lundqvist CB, Borger MA, et al. ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. Eur Heart J. 2012 Oct 1;33(20):2569-619.
 - 16. Recommendations for use of selective and nonselective nonsteroidal antiinflammatory drugs: An American
 College of Rheumatology white paper. Arthritis Rheum. 2008 Aug 15;59(8):1058–73.

BMJ Open

- 17. Tendera M, Aboyans V, Bartelink M-L, Baumgartner I, Clement D, Collet J-P, et al. ESC Guidelines on the diagnosis and treatment of peripheral artery diseases: Document covering atherosclerotic disease of extracranial carotid and vertebral, mesenteric, renal, upper and lower extremity arteries * The Task Force on the Diagnosis and Treat. Eur Heart J. 2011 Nov 2;32(22):2851–906.
 - 18. Camm AJ, Lip GYH, De Caterina R, Savelieva I, Atar D, Hohnloser SH, et al. 2012 focused update of the ESC Guidelines for the management of atrial fibrillation. Eur Heart J. 2012 Nov 1;33(21):2719-47.
 - 19. Smolen JS, Landewé R, Breedveld FC, Buch M, Burmester G, Dougados M, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2013 update. Ann Rheum Dis. 2014 Mar;73(3):492–509.

<text>

Additional file 5 Geographic variation of the health care services grouped by strength and direction of recommendations, and service category



A Weak and strong recommendations; B Positive and negative recommendations; C Diagnostic and treatment services. MOR – median odds ratio. Boxplots depict the interquartile range of values (upper and lower hinges), and the median value.

Based on Welch's t-test, the difference in mean variances [95CI%] of diagnostic and treatment services was 0.04 [-0.01, 0.11], and the difference in mean MOR was 0.11 [-0.01, 0.23].

Page 37 of 37

		BMJ Open <u>50 bn</u> open				
STROBE 2007 (v4) Statement—Checklist of items that should be included in reports of <i>cross-sectional studies</i>						
Section/Topic	ltem #	Recommendation 9	Reported on page			
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1,2			
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2			
Introduction		5 21.				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4			
Objectives	3	State specific objectives, including any prespecified hypotheses	4,5			
Methods		de d				
Study design	4	Present key elements of study design early in the paper	2,			
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	5-7			
Participants	6	(<i>a</i>) Give the eligibility criteria, and the sources and methods of selection of participants	7, Supplementary file 1			
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	7, Supplementary file 1			
Data sources/	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe	6, Supplementary			
measurement		comparability of assessment methods if there is more than one group () / 호	files 1-4			
Bias	9	Describe any efforts to address potential sources of bias	6,7			
Study size	10	Explain how the study size was arrived at	5			
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which grougings were chosen and why	7			
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	7			
		(b) Describe any methods used to examine subgroups and interactions	7			
		(b) Describe any methods used to examine subgroups and interactions \vec{c} (c) Explain how missing data were addressed \vec{c}	NA			
		(d) If applicable, describe analytical methods taking account of sampling strategy	NA			
		(e) Describe any sensitivity analyses Solution	NA			
Results						

 6/bmjopen-2

		02	
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examine of or eligibility,	8,9
		confirmed eligible, included in the study, completing follow-up, and analysed	
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	NA
Descriptive data 14*	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9
		(b) Indicate number of participants with missing data for each variable of interest	NA
Outcome data	15*	Report numbers of outcome events or summary measures	9
Main results 16	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision $\frac{4}{3}$ eg, 95% confidence	9, Supplementary
		interval). Make clear which confounders were adjusted for and why they were included	file 1
		(b) Report category boundaries when continuous variables were categorized	NA
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	11
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	13,14
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	14,15
Generalisability	21	Discuss the generalisability (external validity) of the study results	14
Other information			
Funding 2	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on	15
		which the present article is based	

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in controls in case-control studies.

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 are http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

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