

# Covert checks by standardised patients of general practitioners' delivery of new periodic health examinations: clustered cross-sectional study from a consumer organisation

Franz Piribauer,<sup>1</sup> Kylie Thaler,<sup>2</sup> Mark F Harris<sup>3</sup>

**To cite:** Piribauer F, Thaler K, Harris MF. Covert checks by standardised patients of general practitioners' delivery of new periodic health examinations: clustered cross-sectional study from a consumer organisation. *BMJ Open* 2012;**2**:e000744. doi:10.1136/bmjopen-2011-000744

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2011-000744>).

Received 19 January 2012  
Accepted 2 July 2012

This final article is available for use under the terms of the Creative Commons Attribution Non-Commercial 2.0 Licence; see <http://bmjopen.bmj.com>

<sup>1</sup>International Screening Committee for Austria, Austrian Public Health Association, Vienna, Austria

<sup>2</sup>Department for Clinical Epidemiology and Evidence-based Medicine, Danube University Krems, Austria

<sup>3</sup>Centre for Primary Health Care and Equity, UNSW Sydney, Australia

**Correspondence to**  
Dr Franz Piribauer;  
[franz.p@zaeg.at](mailto:franz.p@zaeg.at)

## ABSTRACT

**Objective:** To assess if data collected by a consumer organisation are valid for a health service research study on physicians' performance in preventive care. To report first results of the analysis of physicians performance like consultation time and guideline adherence in history taking.

**Design:** Secondary data analysis of a clustered cross-sectional direct observation survey.

**Setting:** General practitioners (GPs) in Vienna, Austria, visited unannounced by mystery shoppers (incognito standardised patients (ISPs)).

**Participants:** 21 randomly selected GPs were visited by two different ISPs each. 40 observation protocols were realised.

**Main outcome measures:** Robustness of sampling and data collection by the consumer organisation. GPs consultation and waiting times, guideline adherence in history taking.

**Results:** The double stratified random sampling method was robust and representative for the private and contracted GPs mix of Vienna. The clinical scenarios presented by the ISPs were valid and believable, and no GP realised the ISPs were not genuine patients. The average consultation time was 46 min (95% CI 37 to 54 min). Waiting times differed more than consultation times between private and contracted GPs. No differences between private and contracted GPs in terms of adherence to the evidence-based guidelines regarding history taking including questions regarding alcohol use were found. According to the analysis, 20% of the GPs took a perfect history (95% CI 9% to 39%).

**Conclusions:** The analysis of secondary data collected by a consumer organisation was a valid method for drawing conclusions about GPs preventive practice. Initial results, like consultation times longer than anticipated, and the moderate quality of history taking encourage continuing the analysis on available clinical data.

## ARTICLE SUMMARY

### Article focus

- Can data from a consumer organisation be useful and valid for secondary analysis in health services research?
- Do GPs follow the guideline for preventive service history taking?
- Was the well-recognised time barrier for delivering preventive services also seen in this study in Vienna, Austria?

### Key messages

- Consumer organisation's assessment of GP performance was valid, representative and precise.
- Around one-fourth of GPs failed to achieve the standard for history taking in the new periodic health examination.
- Consultation time was longer than expected and sufficient: time-barrier problem has been overcome.

### Strengths and limitations of this study

- Forty visits at 21 GPs are a small sample; however, this size is comparable to similar mystery patient studies.
- All ISPs went undetected, in contrast to many similar studies.
- The random sample was found to be double stratified and well balanced.
- Multilevel analysis was possible and indicated the role of GP practice style.
- Additional to direct observation data, copies of GPs' record notes may provide further objective assessment.

## BACKGROUND

For many eligible patients the provision of adequate preventive care is blocked by well-known barriers, despite the existence of elaborate guidelines based on best evidence.<sup>1–3</sup> Lack of time and inadequate reimbursement were the main barriers

named by Canadian family physicians to performing the periodic health examination (PHE) as recommended by the Canadian Task force on the PHE.<sup>4 5</sup>

Our main research question was whether the secondary analysis of routine data from consumer associations was feasible to observe quality aspects of the delivery of preventive care. We have not identified any other studies using consumer organisation data for secondary analysis in preventive healthcare performance assessment. As consumer associations with long traditions exist in all industrialised nations, such as Consumer Reports in the USA, similar data could well be available in many countries and could be analysed by health service researchers in the way we propose in this paper.<sup>6</sup>

Studies of preventive service provision which rely on electronic medical record audit, physician self-report, patient surveys and chart review are all prone to bias, as they usually lack validation against observed practice. Studies with standardised patients (SPs) have been used successfully to overcome these kinds of bias.<sup>1 7</sup> An SP is a healthy subject who is trained to assess the performance of doctors based on predefined criteria. Unannounced or incognito standardised patients (ISPs) have been used unobtrusively to assess the routine practice performance of doctors.<sup>8</sup> *"Unknown to the prospective provider of care, such a 'patient' arrives at the clinic and requests care. What happens is gleaned from the records of care and also from the observations reported by the pseudo patients, who have been trained to make the needed observations."*<sup>9</sup>

These ISPs are the healthcare version of the mystery shoppers used in other industries. *"Mystery shopper or visitor are a well known and widely used standardized method in quality management for assessing service quality in the retailing and tourism industry."*<sup>10</sup> Observing health service providers routine or students practical performance by ISPs is a method established since decades in health services and health education research.<sup>11–18</sup> Collecting data by observing performance enables researchers to judge if guidelines are followed, like it has been demonstrated for community pharmacies recently.<sup>19–21</sup> For instance, in the case of PHE delivered by general practitioners (GPs), it could be observed if they ask their patients on their smoking status, as recommended by the preventive service guideline.

In autumn 2008, the official consumer information association of Austria, 'Verein für Konsumenteninformation' (VKI), published a test report on physicians delivering the PHE. In the spring of 2008, two ISPs, members of the VKI tester team, had visited unannounced a sample of randomly selected GPs in Vienna, Austria.<sup>22</sup>

In Austria since 1974, GPs have been reimbursed for annual PHEs from public funds, currently at around US \$100 (€75, current value) per patient. This service is provided free of charge to patients. A reform of the content and new documentation standards were introduced in 2005. Since then, the PHE is based on

a published evidence-based guideline. The evidence base is derived mostly from the USA, Canadian and Australian preventive service guidelines with local adaptations. These guidelines demonstrated by the use of best evidence the causal link of interventions and beneficial medical outcomes for a long list of conditions.<sup>23–26</sup> The interventions recommended in the local Austrian guideline should yield beneficial medical outcomes when performed according to the guideline by GPs. These beneficial screening interventions include, to name a few, assessing smoking status, blood pressure, body mass index, cardiovascular risk calculation and recommended follow-ups like brief smoking cessation advice, etc. Not performing those during the PHE may harm the still healthy patient (client).<sup>27 28</sup> Each year around 850 000 PHEs are performed among the adult Austrian population of 6 million.<sup>22 29–31</sup>

Two types of insurance funding exist in Austria for GPs offering the PHE free of charge. A GP may hold a comprehensive insurance contract plus a PHE contract or a PHE contract only. In our study, we referred to GPs with the comprehensive plus PHE contract as 'contracted GPs' (in Austrian–German '§2 Kassenärzte'), and those with the PHE contract only as 'private GPs' (in Austrian–German 'Wahlärzte mit Vorsorgeuntersuchungsvertrag'). Payment of 'private Austrian GPs' can involve out-of-pocket payments of patients to cover part or all the patient expenses and refunding of a part by insurance. According to a previous study in Austria, the reasons for choosing such a private GP ('Wahlarzt') include short waiting and longer available consultation times.<sup>32</sup> A description of the Austrian health system with its mixed contracted and non-contracted private GP primary care system is beyond the scope of this paper and can be found in an English/German WHO country report.<sup>29</sup> In this study, all GPs had a PHE contract and thus no out-of-pocket payments for any PHE service were necessary, even for 'private GPs'.

The first research question for our secondary analysis was: Did Austrian GPs spend sufficient time to conduct the preventive activities required? Furthermore, we wanted to examine if there was a difference between 'private' and 'contracted' GPs in three quality aspects of care delivered: consultation and waiting time and guideline adherence.

## METHODS

Our methods were structured in two step-like parts. In the first step, we critically appraised the methods used by VKI: their sampling and data collection. In the second, we performed our own analysis on the electronic data set provided by VKI.

Our study design was presented to the legally relevant public health ethics commission of Vienna, which had no objections: the secondary use of these anonymous data on physician performance did not infringe on rights of patients nor physicians.

The GPs' legal representative, the Vienna medical chamber, had agreed at the end of 2007 that some randomly selected GPs may be tested for their PHE performance by ISP from VKI in the upcoming weeks. All GPs of Vienna were informed by their legal representative, about the possible random sample visits. There was no possibility for GPs to opt out.

### Appraisal of VKI sampling and data quality

Knowledge about the VKI methodology was gained through one personal and two phone interviews at the end of 2008 and in first quarter of 2009 with the researcher at VKI who managed the study.<sup>22</sup> We further analysed the note-taking forms used by the ISPs, the VKI's internal written interpretation guide and a report on the VKI testing methodology published in 2008.<sup>33</sup>

We judged the quality of the sample by comparing it with the GP distribution in Vienna and by repeating the VKI sampling procedure in a simulation of our own. We assessed the quality of the data gathered by the ISPs against criteria for a good quality ISP study provided by a recent systematic literature review in the field.<sup>8</sup> These criteria cover the use of content checklists, note taking by the ISP, soundness of clinical cases and ISP detection rates. The results of our appraisal are presented in our first set of findings below.

### Secondary analysis

#### Data preparation

VKI provided a de-identified electronic data set (42 records). In this data set, GPs' names and office locations were deleted and GPs were sequentially numbered by VKI. We transformed the VKI ratings into corresponding numerical values (eg, the five Likert scale satisfaction scores ranging from '+' (very good), through 'o' (average) to '–' (not satisfactory) were re-coded by us into the five integers from 4 to 0). Continuous variables, such as waiting times and consultation

times, were transferred unchanged into our final secondary data set.

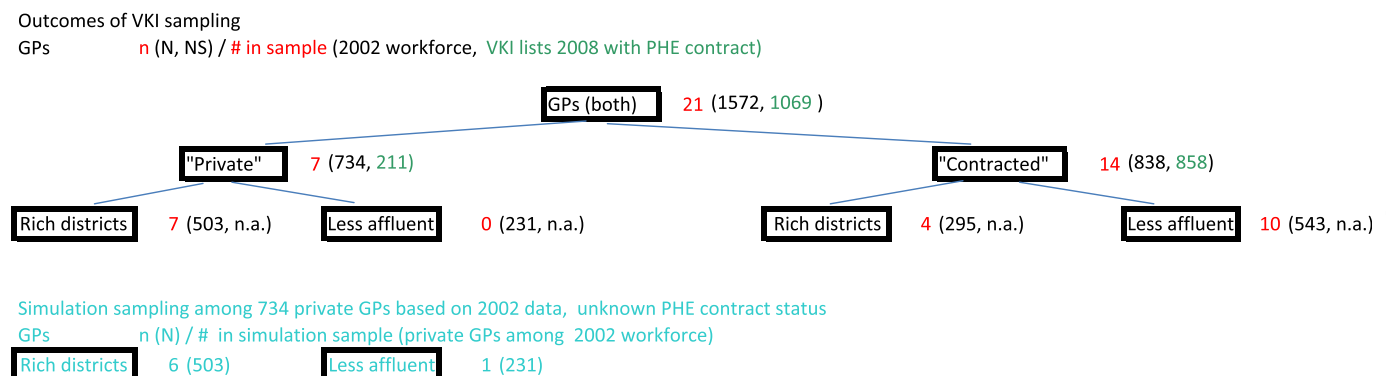
Additionally, we were provided with hard copy clinical results which had been given to the ISPs by the GPs and which were not used by VKI in its own report (34 records—eight were missing). These 34 forms were copies of the double page health summary sheets (HSS, 'Befundblatt'), which the GPs should provide in hard copy at the end of the PHE to their clients.<sup>34 35</sup> One of us (KT) blinded to the medical content of the ISP clinical cases, extracted and coded all clinical data from the 34 paper forms into a second electronic data set in December 2008. More than 90 variables were coded from these data. Free text remarks by the physicians were not extracted (see additional file 1: scanned HSS coding template with data of GP No. 1).

### Statistical analysis

We found a double stratified probabilistic sampling. GPs were drawn by VKI within their two strata, private/contracted (stratum 1) and district blocks (stratum 2) by a strictly random process.

The primary sampling unit for our data analysis was the GP (see figure 1). Each of 21 practitioners were offered a visit by the two different ISPs. Two of the practice visits were rejected by two GPs—one private and one contracted (because of an administrative error and because laboratory results were not ordered by the GP). Both GPs were visited by the other ISP. The visits resulted in a total of 40 observations on 21 GPs, belonging either to the 'private' or to the 'contracted' GP group. The clustering at the GP was accounted for in our statistical analysis by survey/panel data methods and additionally by multilevel data analysis.<sup>36</sup> The reasons for the multilevel analysis are explained below in the appraisal of sampling by VKI.

We conducted our statistical analysis for this publication with Stata V.9.1 and 11.<sup>37</sup> Descriptive statistics (eg, means,



**Figure 1** Results of Verein für Konsumenteninformation (VKI) sampling compared with our simulation sampling of private general practitioners (GPs). In 2008, 21 GPs were sampled by VKI, seven of them "private GPs". All seven were located in the richer part of Vienna. Among the "contracted GPs", four of 14 were located in the richer Vienna districts. GP workforce data of 2002, published in a health report of the city of Vienna administration, provided the most recent information on distribution of private GPs among the Vienna city districts. As we were not provided with data, beyond totals, on the two sampling population lists of VKI. n.a., could not access the district distribution data; PHE, periodic health examination.

proportions and CIs) were produced by the Stata survey/panel data methods with the most conservative assumptions (eg, finite-population assumption, linearised proportions and binomial Wald statistics for CI of proportions). For additional modelling, we used mixed-effects restricted maximum likelihood estimation and generalised linear models for continuous variables and random- or fixed-effects logistic regression for binary dependent variables (multilevel data modelling). All statistical tests performed and CIs reported are at the 95% level.

For performance assessment, we constructed appropriate indicator variables in accordance with the published guidelines for the PHE based on the observations of the ISPs.<sup>23</sup> For example, only if the full structured medical history proforma, the 'health information sheet' (HIS), was completed, including optimal alcohol screening according to guideline, was the constructed binary (yes/no) indicator coded positively.

## RESULTS

### Step 1: appraisal of VKI sampling and data quality

#### Sampling GPs

VKI reported to us that they used a double stratified random sampling method for GPs in Vienna. One stratum was insurance contract status ('private/contracted') and the other was the geographic distribution of doctors among 23 districts in Vienna. Two independent numbered name lists, one for 'private GPs' and another for 'contracted GPs', were used. The lists were provided to VKI by the Central Association of Austrian health insurances ('Hauptverband der österreichischen Sozialversicherungsträger'), which runs the central registry of all PHE contracts but not to us. Each list was sorted for districts, showing the office locations and the total number of GPs in each district. The sample population in the lists was 1069 GPs, 211 (20%) of whom were 'private'. VKI fixed the GP sample size at 21, seven of whom (33%) being 'private GPs', thus creating deliberately an oversampling of 'private GPs' as they explained in the initial interview.

To determine the sample size per district block, the number of GPs to be sampled for each district was calculated by VKI from the names lists sorted for districts. For example, the seven 'private GPs' were sampled from a workforce distributed over 23 districts. Each of the seven district sampling blocks formed should comprise around 14% of the workforce. Thus, districts were lumped together in the sorted list until a block held around 14% of the 'private GPs' workforce, then the next block was created from the remaining districts and so on. In this way, the number of GPs per district was fixed for all 23 districts in Vienna and for each of the two GP contract types separately.

Selection from a district block was done by drawing a random number within the numbered name lists. The random number for each district block was generated by an internet-based public domain software, AGITOS. The sampling base numbers used in AGITOS for each block

was determined by the total number of GPs in each district block.<sup>38</sup>

After the GPs' names were determined, the ISPs arranged the visits. If an appointment could not be arranged, the ISP called the VKI office and a replacement GP was drawn there by the random number mechanism within the district block, as described above. To visit seven 'private GPs', 14 replacements were needed. This contrasted with three replacements needed for the 14 'contracted GPs'.

The VKI methodology resulted in one GP being selected in 15 of 23 districts, two GPs in three districts (Nos. 3, 18 and 19) and no GPs in five districts (Nos. 5–8 and 17) (see table 1). Six GPs in the sample were from inner districts and 15 from outer districts. Eleven GPs had their office in the more affluent part of Vienna and 10 in the less affluent. The nine inner city districts (Nos. 1–9) in combination with three outer districts (Nos. 13, 18 and 19) comprised the more affluent part of Vienna compared with the rest, judged by purchasing power per head and housing prices (for details classifying affluent

**Table 1** Outcome of VKI sampling of GPs in Vienna by city district and GP insurance contract

Vienna district, No.	VKI sample (number of GPs)	Of these	
		Private*	Contracted†
1	1	1	
2	1		1
3	2	1	1
4	1	1	
5			
6			
7			
8			
9	1	1	
10	1		1
11	1		1
12	1		1
13	1	1	
14	1		1
15	1		1
16	1		1
17			
18	2	1	1
19	2	1	1
20	1		1
21	1		1
22	1		1
23	1		1
Total GPs	21	7	14

VKI published the names of the GPs sampled and their office locations in its report October 2008.

\*We assessed the contract status of each named GP through the public internet search template of the Vienna Medical Chamber (<http://www.praxisplan.at>).

†'Contracted GPs' have a full contract with the regional general health insurance including a periodic health examination contract ('S2 Kassenärzte').

GPs, general practitioners; VKI, Verein für Konsumenteninformation.



versus less affluent districts, see additional file 3: GP sample distribution in rich and poor parts of Vienna).

The distribution of sampled GPs among the Viennese districts should resemble as much as possible the distribution of the real GP workforce performing PHE among the districts. The stratification aimed to improve the representativeness with regard to two strata, geographic distribution and insurance contract status. 'Contracted GPs' per district should correlate with the district population size, as 'contracted GPs' are placed by the Vienna general social insurance agency to serve the population. Thus, highly populated districts should also be represented well in this sample. Inner city districts (Nos. 1–9) have a smaller population than most of the 13 outer ones (Nos. 10–23). The sample reflected this distribution, with a GP ratio of 6:15 for inner versus outer districts. 'Private GPs', meanwhile, are free to establish themselves wherever they like. We assumed that they would tend to open their offices in the more affluent districts, as their income relies on out-of-pocket payments for most of their services except the publicly financed PHE.

To examine the quality of the random sample block procedure of VKI, we had to rely on other data, as we were not given access to the two original VKI sampling population GP lists. Only the totals of their two lists were reported to us, namely 211 'private GPs' and 858 'contracted GPs'. We repeated and thus simulated the VKI procedure with the most recent and applicable data we could find. These were published by the city administration of Vienna in 2002, reporting on the district distribution of 734 private GPs out of total of 1572 GPs.<sup>39 40</sup> Data on PHE contracts of these private GPs were not available. According to that data many of the private GPs (17%) practised in the 19th (9%) and 13th (8%) districts. When repeating the VKI's district block procedure with this other data, the first of the seven GPs was drawn by us out of the first block composed of those two districts. The next two (1st and 18th) did hold together 14%, so the next GP was drawn from this second bloc and so on. In our simulation, the seventh 'private GP' was drawn from five districts at the end of the list, each with <3% of the workforce (see also additional Excel file 4: sampling assessment including source data and further 2007 city administration workforce data).

When comparing our simulation result with the sampling result of VKI, published in its magazine with GP name and location, we found a nearly identical distribution.<sup>22</sup> In the VKI sample, all seven 'private GPs' were from the rich part of Vienna, whereas in our simulation, six of the seven were from that part. However, as only 211 'private GPs' held a PHE insurance contract in 2008, the district distribution of 211 'private GPs' in the VKI list might be different from that of the 734 private GPs of our data of 2002. This could explain the small deviation from our simulation result (see figure 1).

### VKI sampling supports level analysis

VKI used a double stratified probabilistic sampling. One stratum was 'private' or 'contracted' GPs. The other strata were the 23 district blocks as described above. By such an intensive stratification and a strictly random selection out of these strata, VKI achieved in our opinion a well-balanced and representative random sample of the GP workforce in Vienna, despite the small sample size of 21 GPs.

After judging the sampling process robust enough, we sought for the most appropriate type of analysis of these data. The two observations dealing with one GP were not independent and thus were 'clustered at the level of the GP'.

We adjusted for this by two types of analysis: correcting for the clustering effect and using multilevel modelling. By multilevel modelling, we could also estimate intraclass effects at the GP level, as proposed in the literature.<sup>36 41</sup>

### Validity of clinical cases

Two ISP clinical cases were constructed by VKI health experts on the basis of the Austrian PHE guideline, available in print and internet download since 2005.<sup>23</sup> The guideline was intended to be used by health service administrators (such as screening programme managers at local and regional levels) to organise the preventive service activities of GPs in their area, similar to guidelines by other professional bodies.<sup>26 42</sup> With the support of medical journalists, the guideline was written to be understandable to a broader audience than GPs, although it includes evidence-based references.<sup>43</sup> The high amount of detail in the guideline allowed VKI experts to develop the two clinical cases for the ISPs in such a manner as to elicit clearly observable actions by the GPs during the PHE.

Both the male and the female ISPs were over 65 and presented complex clinical screening cases. The predominant critical screening task of the male was the detection of his high cardiovascular risk and of the female her clearly problematic alcohol consumption. However, the task involved screening for nearly all 15 target conditions of the Austrian PHE.

Apart from the clinical case history, the two ISPs presented the GP with fabricated laboratory data, tailored to their cases. For example, the woman reporting problematic alcohol consumption had elevated levels of serum liver enzymes (Gamma GT: 65 U/l, GOT 44 U/l, GPT 35 U/l). Before the fieldwork, the ISPs rehearsed with the help of the outpatient facility of the Vienna public social insurance medical service, where also their laboratory details were fine tuned. A more detailed description of the clinical case construction is included as additional file (see additional file 2—'ISP\_Cases').

### Assessment of data collection by ISPs

The two ISPs each arranged visits with 21 GPs. At the GP's office, each ISP completed the standardised

evidence-based HIS, a questionnaire which all GPs offering reimbursed PHE are obliged to provide.<sup>44</sup> They also completed the AUDIT-GMAT, an Austrian version of the WHO questionnaire 'AUDIT' for problematic alcohol consumption, when offered.<sup>45</sup> The ISP training had included completion of the HIS and AUDIT-GMAT as well as presentation of their history personally to the GP. At the end of the consultation, they each collected the standardised HSS, which the doctor is also obliged to complete and provide in copy to his/her client. More information about the standardised medical records set for the Austrian PHE is detailed below in the results and has been published elsewhere.<sup>46</sup>

Immediately after having left a GP's office, the ISPs noted their experience using a standardised note form. At the VKI office, an independent person extracted data for the calculation of scores. The data coding was explicitly defined for the GP test in advance by specifically written instructions called 'Regeln für die Eingabe/Beurteilung in TestRev' (rules for data entry and assessment into TestRev). We were provided with these specific coding rules. TestRev is the routine software and database VKI applies for storing, analysing and reporting on the numerous tests they perform in all fields of industry and services. For data handling, an in-house quality management handbook exists, and this was also applied for the PHE test. VKI holds an official state quality certificate for its testing procedure.<sup>33</sup> After data entry, a second person compared the extracted results in TestRev with the protocol notes of the ISP. In the case of disagreement, a third independent senior person decided as to the correct interpretation and coding.

In this way, VKI gathered in its electronic data set detailed and summary statements such as the ISPs' subjective impressions (satisfaction), but mostly VKI gathered more objective observations on activities the GPs performed or omitted. These more objective ISP observations can be considered in the healthcare quality field as 'patient experience', more amenable to effectively improving quality of care than the more subjective 'patient satisfaction'.<sup>47–50</sup> VKI condensed the ISP notes into 45 statements/judgements per visit. This 45 items VKI data set was made available to us. We were not provided with the notes taken by the ISPs. However, as the strict rule-based coding system of VKI allows the condensed statements/judgements to be re-expand to the detailed observations, we could interpret the performance of each GP to a greater degree than the 45 items would suggest. For example, problematic alcohol consumption should be screened for. VKI coded '+ +' (very satisfactory) when the AUDIT-GMAT questionnaire was handed over to the ISP, 'o' (average) when the questionnaire was not used but the GP did discuss alcohol consumption with the ISP and '– –' (not satisfactory) when the topic was not even raised verbally.

We found the VKI method to be reliable in reporting on the ISPs' experience of GP interventions, which should have been performed during the PHE. For this

first publication, we restricted ourselves to analysing data on waiting and consultation time, and GP performance during the medical history taking phase, compared with guideline recommendations.

### Detection rate of ISP

Detection of ISP by the observed physician can be an important obstacle in ISP studies,<sup>8</sup> leading to bias and confounding. We are confident that all ISP visits went undetected and physician behaviour was not distorted by the idea that the client could be an expert observer with a constructed clinical case. The age of both ISPs was the same as in the presented clinical cases. Great care was taken to ensure that there was no observable difference on signs. The responsible researcher at VKI stressed in the first interview with us in October 2008 that none of the 40 ISP visits had been detected. We asked her again in February 2009 to interview the two ISPs to determine if they had any suspicion that any of the GPs could have detected them. The response was again negative. One ISP even replied on that occasion that the only GP who had seemed to be a little suspicious had just sent a personal invitation letter to return for the next annual PHE.

### Results of step 2: secondary analysis

In our secondary analysis, we focused primarily on observational experience data. The satisfaction data have been published by VKI in its own magazine.<sup>22</sup> We received data on 40 of 42 arranged ISP visits, the same number as reported in the VKI test report publication in 2008. Two ISP visits were rejected by two GPs, one 'private' and the other 'contracted'. The reasons given by the two GPs for rejection were in one case an administrative GP error (a misunderstanding of the use of the electronic insurance patient access card) and in the other that the pre-prepared laboratory results were not ordered by the GP herself. However, both GPs were visited by the other ISP.

### Service delivery time

For the completed visits, the average consultation time was 46 min (95% CI 37 to 54 min). For the male ISP, it was 38 min (95% CI 33 to 43) and for the female ISP 54 min (95% CI 40 to 67). The difference of 16 min between the two ISP cases was not significant, when applying a survey/panel data method adjusting for the clustering effect at GP level, but was significant in the full-adjusted multilevel model (coefficient 15.6, 95% CI 4.9 to 26.3, see figure 2).

Female GPs offered longer consultations, with an average of 47 min (95% CI 38 to 57) than males, with an average of 38 min (95% CI 19 to 58). The observed difference of 11 min in our sample is not significant, when applying the survey/panel method adjustment for multilevel modelling or the full-adjusted model (see figure 2).

Using multilevel analysis, we estimated the proportion of variance explained by the intraclass effect versus the

**Figure 2** The full-adjusted multilevel model (generalised linear model, restricted maximum likelihood).

$$\text{Var (Dep. Variable (eg, waiting time))} = \theta_0 + \theta_1(\text{GP}) + \theta_2(\text{GP gender}) + \theta_3(\text{GP contract type}) + \theta_4(\text{ISP case})$$

difference between the GPs. If a high proportion of variance is explained by one variable, then this variable has a strong effect on the outcome of interest. Sixty-two per cent of the variance for waiting time was determined by the GP intraclass effect compared with 30% for consultation time. These variance estimates result from a conservative monovariate random effect generalised least squares regression model with the GPs as explanatory variable. Further adjusting for the two different ISP case types increased the variance proportion for consultation time explained by the GP by one-third, to 45%. The same adjustment did not significantly change the variance proportion in waiting time (slightly increased from 62% to 67%). As could be expected, the intraclass and adjustment effects were even more pronounced in the fixed random effect model (see table 2).

We also found a difference of 22 min in average consultation time between private and contracted GPs. The difference was significant. 'Private GPs' provided 60 min (95% CI 50 to 71) and 'contracted GPs' 38 min (95% CI 26 to 49) on average. The difference remained significant using a fully adjusted multivariate model, which included the two ISP case types, GP gender, GP insurance type and the clustering on the GP level (generalised linear modelling statistics incorporated in Stata V.11.0) (see also additional file 9: STATA-Commands (selected).txt)

### Quality of service

For this publication, we compared observed GP history taking performance with the evidence-based recommendations. According to the officially published guideline, the PHE should include a structured general history taking supported by the HIS and questions regarding alcohol use supported by the AUDIT-GMAT.

We classified five performance levels in respect to general history taking adherence to the guideline before analysing the data. The five HIS scores ranged from '0' (= below minimal) to '4' (= perfect history). The maximum general HIS score of 4 was achieved when the HIS was offered and all medical domains were addressed during the consultation. Omission of one of the eight medical domains was tolerated in our data interpretation as possible measurement error on the part of VKI. A score of '3' was achieved when the HIS was offered but not all domains were touched on additionally verbally. No HIS, but raising at least seven of the eight required domains verbally scored '2'. A score of '1' was given when there was no HIS and two or three domains were missing. No HIS and four to eight domains not addressed scored '0'. As the general PHE contract with the GPs requires that the HIS proforma be completed, we considered HIS scores of '2' or less below standard.<sup>51 52</sup>

Screening for problematic alcohol consumption should start with completion of the AUDIT-GMAT questionnaire by the client. For this screening activity, we scored the performance into two categories. Care according to guideline provided the AUDIT-GMAT (we scored '1'), otherwise we scored '0'.

A HIS was offered in 53% (95% CI 34% to 71%) of all visits. Among the GPs offering a HIS, a proportion outperformed the requirements of the guideline if they additionally addressed nearly all the medical content of the HIS during the consultation phase of the PHE (HIS score '4'). In 20% of all visits, GPs scored '4', indicating perfect general medical history taking (95% CI 9% to 39%).

The AUDIT-GMAT was offered in 38% (95% CI 19% to 56%) of all visits. There was no difference between 'private' and 'contracted GPs' ( $p=0.89$ ) and no

**Table 2** Proportion of all variance explained by intraclass (GP) variation in multilevel analysis on waiting and consultation time

Time	Random effect (conservative)	Fixed effect (strong assumption)
Regressed on GP only		
Waiting	0.621*	0.686
Consultation	0.298	0.493
Time	Random effect	Fixed effect
Regressed on GP and ISP (adjusted for ISP case type)		
Waiting	0.668	0.718
Consultation	0.445	0.562

After adjustment for ISP case type, the intraclass effect of waiting time did increase a little, whereas for consultation time, the effect increased from 0.30 to 0.45 in the random effect model. The conservative random effect model seems to us most appropriate for this kind of data, especially due to our small sample size.

\*rho: proportion of all variance explained by intraclass (GP) variation.  
GPs, general practitioners; ISP, incognito standardised patient.



difference between the female and male ISPs ( $p=0.73$ ). All GPs who offered an AUDIT-GMAT had also offered a HIS (see also additional file 5: HIS and AUDIT scores crosstable  $n=40$  cases).

We considered the acceptable overall history taking service standard level to be a HIS offered (HIS score '3' or higher) plus the alcohol topic addressed at least verbally. Thirty per cent (95% CI 12% to 48%) of all visits were performed below this standard. The difference in proportion of 'private GPs' (21%) and 'contracted GPs' (35%) was not significant in the full multilevel model ( $p>0.05$ ).

We found a significant intraclass effect at the GP level: for a given GP, the OR was 60% (95% CI 0.03% to 91%) that their consecutive next ISP would also get the same level of medical history performance. This intraclass effect indicates that GP practice style was a determinant of history taking performance.

## DISCUSSION

Our study is the first using direct observation via ISPs of routine preventive service GP performance compared with standards in an evidence-based structured national PHE programme. We have been unable to find any similar previous studies, which used secondary data collected by mystery patients, ISPs, engaged by a consumer organisation. The Austrian consumer organisation (VKI) evaluated GPs' performance in Vienna in delivering preventive care, specifically the highly standardised Austrian PHE for which a curtailed evidence-based guideline is published in German since 2005.<sup>23</sup> The random sampling process for GPs appears to have been sound and produced a representative sample. The clinical cases for the ISPs fitted well to the physical appearance of the two ISPs, one male and one female around 65 of age. In none of the 40 completed visits was there any evidence that the ISP had been detected by the GP. The 40 cases were clustered at the level of 21 GPs. The GP sample had two strata. The first strata were 'contracted GPs' and 'private GPs'. The 'private GPs' were slightly oversampled (by three GPs) as their proportion was 33% in the sample and 20% in the sampling population of 1069 GPs with PHE contract in 2008.

The second level, Vienna city districts, improved the sampling quality further, as the random sampling procedure within the city district blocks was found to be robust. Generalisation of the findings to the Viennese GP workforce delivering the PHE is reasonable within the statistical limits of the small sample.

## Limitations and strength

One limitation of our study is the small sample size of 40 completed ISP cases for 21 GPs in the VKI data set. In a recent systematic literature review of good quality SP studies by Rethans,<sup>8</sup> a median 39 GPs were visited across the 20 studies reporting on GPs since 1985. There has been a trend to smaller studies since 2000, with a median of 27 GPs. Our small sample size means that the

estimates have wide CIs, especially when considering subgroups, such as 'private GPs'. Only when effect sizes are large, for example, in our case, when expected values differ dramatically from observed ones, can we rule out chance.

Measurement error on the part of the ISPs is an important potential threat to validity. Rethans proposes that this can be overcome by thorough ISP training, case preparation and robust documentation processes. In the VKI study, the two ISPs were highly experienced, having worked more than 2 decades in consumer testing of many service industries. The VKI tests run now in the thousands—the test of the Viennese GPs on the PHE is just one of the assessments they have performed. More than 80 tests are conducted each year, the organisation has existed for more than 3 decades and is internationally recognised among European consumer organisations. It has an ISO quality certificate for its testing procedures and constant internal quality checks. The data have to be well documented and robust, as legal cases are common, with tested providers or producers often appealing to the courts.<sup>33</sup> In summary, our primary data collection was embedded in a high-volume routine with sound quality assurance and collected by highly trained professionals, and thus, the data are likely to be reliable.

The data collectors themselves (ISPs) were blinded to our (implicit) study hypotheses, such as expected duration of consultations being 5–10 min. It could be argued that consumer associations may be especially critical of doctors and that this might have affected the study design and data collection. In this case, however, the Austrian VKI test report signalled satisfaction with GPs' PHE performance (translated title: 'PHE in good hands')—in contrast to their repeated critical reports of ISPs observing pharmacies.<sup>22</sup>

However, the satisfaction of VKI can only be a weak proxy for a real satisfaction study, as a further limitation of ISP studies is that they cannot measure an important component of quality outcomes: patient satisfaction. It can only be assessed from real patients, for example, by surveys. In the case of PHE, satisfaction will be important, as satisfied clients tend to return, and follow-up at the recommended screening intervals. Several large surveys, although most probably not representative due to a low response rate of around 30%, have been done by others recently for the new Austrian PHE and signal a satisfaction level of 41% being very satisfied with the quality.<sup>53 54</sup> The measurement of satisfaction levels has its own limitation in international comparability, when self-developed questionnaires are applied locally, as observed satisfaction levels are highly depending on the content and framing of the questions.<sup>55</sup>

Several other important aspects of quality of care, like communication skills of GPs and knowledge of GPs on prevention have not been looked at by the ISP and cannot be addressed in our study.<sup>56</sup>



A strength of our data, in contrast to many other ISP studies, is that all ISP visits were undetected. Furthermore, our study was not distorted by a self-selection bias of voluntarily participating GPs. In other studies, around 40% of physicians on average decline to participate, leading to a severe self-selection bias among physicians.<sup>1 8</sup> We were able to completely avoid this bias by using the anonymous data collected by VKI, as GPs were selected by a strict and sophisticated random sampling procedure. The Viennese Chamber of Physicians agreed collectively to participate, and single GPs could not exempt themselves from the random VKI visits. The visits to few of around 1500 GPs were announced to all by their Viennese medical chamber, without giving an exact date. However, the VKI never asks permission at the individual service provider level.

### 'Lack of time' barrier

One of the main obstacles or barriers named by GPs worldwide to delivering preventive care is the lack of time.<sup>5</sup> Among others factors, administrative arrangements including financial factors are important to consider when routine GP practice needs to be changed.<sup>26 57</sup> The average consultation time of 38 min among the 'contracted GPs' (§2 Kassenarzt) is much longer than the 10–15 min we expected when the PHE reform was set in motion by one of us (FP) in 2003. Austria has a kind of capped fee-for-service system for 'contracted GPs', which results in high volumes of services and high turnover of patients.<sup>29</sup> We estimate the average consultation time to be in the range of Germany with its 7.6 min, found in the most recent comparative, but not representative, study in Europe.<sup>58</sup> No study using representative data has been published in a peer-reviewed journal on this issue for Austria.

The 60 min consultation time with 'private GPs' in this study is extraordinary, especially as these consultations are available free of charge to the eligible population. However, it was difficult for the ISPs to secure an appointment with 'private GPs'—they had to contact 21 to make appointments with 7 (1:3 ratio). Thus, the PHE is a scarce commodity in private practice and its widespread uptake would likely result in waiting lists.

The long average consultation time of 46 min may also be attributable to the complex ISP cases, as increased severity of cases leads to longer consultation all over the world.<sup>59</sup> Less complicated cases, especially among younger clients, would be more the norm and these may be handled in a shorter time. The consultation duration for less complicated cases is unknown and requires further research in Austria.

The Austrian model, developing guidelines accompanied by standardised report cards in combination with a generous reimbursement system based on special contracts for prevention (the PHE contracts), could obviously overcome the barrier of limited time available in Vienna general practice.<sup>46</sup>

### Service quality—times typical for GPs

In addition to the sufficient time spent on average to perform the PHE, we observed intraclass effects at the individual GP level for consultation and waiting times. The GP effect was stronger on waiting time than on consultation time. In other words, each GP tended to have a typical waiting and less so consultation time, being repeated with the second visitor. Such a typical behaviour, which we called in accordance with the quality management literature 'practice style', is thought to formed over longer times by various factors.<sup>60 61</sup> These may be non-GP factors, like patient load per day or usual severity of cases, depending on the area a GP works. We have found 'private GPs' being highly concentrated in the richest districts of Vienna, whereas 'contracted GPs' were distributed according to population per district (see results on sampling above). From the positive correlation of social class and health status in the population follows, that 'contracted GP' tend to have poorer, sicker, less educated patients, as only the well-off can easily afford a 'private GP'. The service of a 'contracted GP' is free, whereas the out-of-pocket payment at the 'private GP' is only refunded to a small part by the health insurances. As all patients are insured in Austria, the richer ones can consume the 'private GPs' in addition to the 'contracted GPs'. The contracted GPs have usually fuller waiting rooms and much more patients per day to serve. The main motive to visit a 'private GP' is to buy and get longer consultation times according to a recent Austrian study.<sup>32</sup>

GP factors like age, training-level and guideline adherence should be typical for the Vienna GP workforce and should not differ among our study subjects systematically, as random sampling should even out those differences. However, the sample was intentionally stratified on contract status of GPs, and 'private GPs' were oversampled by VKI, as the consumer organisation hypothesised a major difference in the delivery of preventive care based on GP contract status.

Income is a further important contributing factor for physician behaviour.<sup>62 63</sup> However, as all 'private GPs' in our sample lack only a general insurance contract, but hold a PHE contract, they do not get any out-of-the-pocket payment for their PHE service. The PHE reimbursement is the same at €75 (around US\$100) for both GP contract types. Thus, the observed tendency of 'private GPs' to counsel longer than 'contracted GPs' cannot be attributed to a direct financial incentive for this service. It seems more to be the 'habit' or patient management style of 'private GPs', which we short named 'practice style' above, as a higher income per case allows 'private GPs' to spent more time per visit.<sup>32</sup>

The results that (1) waiting time was mainly influenced by the GP and (2) consultation time was mainly influenced by the clinical case presented are also congruent with knowledge from quality management on practice styles and results from health services research.<sup>9 58</sup>

In summary, the practice style of GPs had a strong influence on waiting time and a lesser influence on consultation time. Consultation time was dependent on the type of ISP case, but waiting time was not. GPs reacted to the specific cases in adjusting their consultation time.

### Service quality—guideline adherence

Overall, history taking standards were missed by 21% of 'private' versus 34% of 'contracted' GPs. This difference was not significant. Multilevel analysis revealed that performing below standard history taking was consistent at the GP level between the two ISP visits. This finding is a further indication of the existence of GP personal practice styles influencing service quality and indicates an opportunity for improvement through training and feedback.

The use of the standardised assessment of a history of problematic alcohol consumption, the AUDIT-GMAT questionnaire, is highly recommended in the guideline for the PHE.<sup>64</sup> Yet in 2005, there was strong opposition voiced against the routine use of this questionnaire by unionised doctors (medical chamber). They considered the questionnaire to be too intrusive and were concerned that it would discourage potential clients. When in 2003 one of us (FP) led the development team for the new PHE, it was expected that only a minority of GPs would apply the AUDIT-GMAT. However, in this study, it was used in nearly 40% of visits, with no significant difference between 'private' and 'contracted GPs'. Many GPs may consider screening for problematic alcohol consumption to be important in a country like Austria with high alcohol consumption.

### CONCLUSION AND OUTLOOK

Using ISPs is a well-established but complex method for health service research. Using data not designed for research is also complex. However, the increase in complexity is outweighed by the reduced bias from unannounced visits. Our study was the first to report physicians' routine preventive performance under direct observation of experienced ISPs applying standardised quality-assured documentation in a nationwide PHE programme. This study mainly reports research methods and length and variation in consultation times and guideline adherence in regard to alcohol screening and medical history taking. Some better than expected results were found, such as the long consultation times and the relatively high completion rate of the alcohol screening questionnaires. 'Private GPs' and 'contracted GPs' did differ more in waiting time than in consultation time and not in regard to alcohol screening. This leads us to a new hypothesis that there is little relevant difference in the medical quality of the service of 'private' and 'contracted GPs'. Further research on the clinical part of our secondary data should help to clarify this issue. We hope that this paper will stimulate health service research on the quality of service of

annual PHEs provided to many of a national population each year.

**Acknowledgements** We thank Dr. Bärbel Klepp, health service specialist at VKI Vienna, for reading drafts and her comments on the methods section of this paper. We further thank her and the management of VKI for their availability for our interviews on the VKI methods and the provision of the anonymised data used for secondary analysis in our study. We further thank the members of the International Screening Committee for Austria, Reli Mechtler, University of Linz, Austria, Gerald Gartlehner, Centre for Clinical Epidemiology, Danube University Krems, Austria, Claudia Luciak-Donsberger, Medical University of Vienna, who commented on draft versions. Ken Eaton helped us finding titles for the series of our four papers planned. A special thanks goes to Sir Muir Gray, NHS Knowledge Centre, UK, who gave valuable comments and motivated the authors as a member of the International Screening Committee for Austria to continue their work. This study was made possible by using the results of a cooperation project between the Central Association of Austrian Social and Health Insurances ('Hauptverband der österreichischen Sozialversicherungsträger') and VKI producing the PHE test report in the VKI publication September 2008.

**Contributors** FP conceived the study, performed the statistical analysis and drafted the manuscript. KT extracted data, helped in the interpretation and finalisation of the manuscript. MFH helped in the interpretation, internal review and finalisation. All authors read and approved the final manuscript.

**Funding** This research received no specific grant from any funding agency in public, commercial or not-for-profit sectors.

**Competing interests** None.

**Ethics approval** Ethics approval was provided by the Statutory Public Health Ethics Committee of the City of Vienna.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** The data of this study are owned by the Austrian Consumer Organisation (Verein für Konsumenteninformation, VKI). On our written request in October 2008, VKI provided us with the electronic data set (raw data: Excel file, 23 lines), and hardcopies of the completed medical result sheets (34 sheets) for the sole purpose of conducting health service research studies by us, the International Screening Committee for Austria. We extracted data from the hardcopies and added it to our own secondary data set. We encourage any researcher to ask permission and perhaps request the data set also from VKI in Vienna, Austria (<http://www.vki.at>).

### REFERENCES

1. Hutchison B, Woodward CA, Norman GR, *et al.* Provision of preventive care to unannounced standardized patients. *CMAJ* 1998;158:185–93.
2. Yarnall KS, Pollak KI, Ostbye T, *et al.* Primary care: is there enough time for prevention? *Am J Public Health* 2003;93:635–41.
3. Baker R, Camosso SJ, Gillies C, *et al.* Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* 2010;(3):CD005470. <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD005470>
4. Fletcher SW, Spitzer WO. Approach of the Canadian Task Force to the Periodic Health Examination. *Ann Intern Med* 1980;92:253–4.
5. Hudon E, Beaulieu MD, Roberge D. Integration of the recommendations of the Canadian Task Force on Preventive Health Care: obstacles perceived by a group of family physicians. *Fam Pract* 2004;21:11–17.
6. Consumer Reports.Org. *Consumer Reports Health.org*. 2010. <http://www.consumerreports.org/health/home.htm> (accessed 7 Jan 2010).
7. Beaulieu MD, Rivard M, Hudon E, *et al.* Using standardized patients to measure professional performance of physicians. *Int J Qual Health Care* 2003;15:251–9.
8. Rethans JJ, Gorter S, Bokken L, *et al.* Unannounced standardised patients in real practice: a systematic literature review. *Med Educ* 2007;41:537–49.
9. Donabedian A. *An Introduction to Quality Assurance in Health Care*. Oxford: Oxford University Press, 2003.
10. Fache W. Methodologies for innovation and improvement of services in tourism. *Managing Service Quality* 2000;10:346–66.
11. Bokken L, Linssen T, Scherpier A, *et al.* The longitudinal simulated patient program: evaluations by teachers and students and feasibility. *Med Teach* 2008;1–8.

12. Gorter S, Rethans JJ, Van der HD, *et al.* Reproducibility of clinical performance assessment in practice using incognito standardized patients. *Med Educ* 2002;36:827–832.
13. Linssen T, Bokken L, Rethans JJ. Return visits by simulated patients. *Med Educ* 2008;42:536.
14. Rethans JJ, Sturmans F, Drop R, *et al.* Assessment of the performance of general practitioners by the use of standardized (simulated) patients. *Br J Gen Pract* 1991;41:97–99.
15. Rethans JJ, van-Leeuwen Y, Drop R, *et al.* Competence and performance: two different concepts in the assessment of quality of medical care. *Fam Pract* 1990;7:168–74.
16. Rethans JJ, van Boven CP. Simulated patients in general practice: a different look at the consultation. *Br Med J (Clin Res Ed)* 1987;294:809–12.
17. Saebo L, Rethans JJ, Johannessen T, *et al.* [Standardized patients in general practice—a new method for quality assurance in Norway]. *Tidsskr Nor Laegeforen* 1995;115:3117–19.
18. Bray BS, Schwartz CR, Odegard PS, *et al.* Assessment of human patient simulation-based learning. *Am J Pharm Educ* 2011;75:208.
19. Watson MC, Bond CM, Grimshaw J, Johnston M. Factors predicting the guideline compliant supply (or non-supply) of non-prescription medicines in the community pharmacy setting. *Qual Saf Health Care* 2006;15:53–7.
20. Watson MC, Cleland J, Inch J, *et al.* Theory-based communication skills training for medicine counter assistants to improve consultations for non-prescription medicines. *Med Educ* 2007;41:450–9.
21. Lin K, Travlos DV, Wadelin JW, Vlasses PH. Simulation and introductory pharmacy practice experiences. *Am J Pharm Educ* 2011;75:209.
22. Klepp B. *Vorsorgeuntersuchung im Test—In guten Händen.* Konsument 2008[9], 22-25. Wien, Austria, 2008.
23. Piribauer F, Schmid D, Pueringer U, *et al.* *Vorsorgeuntersuchung Neu - Wissenschaftliche Grundlagen.* In: Pueringer U, Ehgartner B, Smolek I, *et al.*, eds. Graz, Austria: Wissenschaftszentrum Gesundheitsförderung/Prävention der VAEB, 2005:1–207.
24. USPSTF. *Guide to Clinical Preventive Services.* 2004<http://www.ahrq.gov/clinic/uspstfix.htm> (accessed 5 Apr 2004).
25. Canadian Task Force on the Periodic Health Examination. *The Canadian Guide to Clinical Preventive Health Care.* 1 edn. Ottawa: Canadian Government Publishing, 1994.
26. Litt J. *Putting Prevention into Practice. Guidelines for the Implementation of Prevention in the General Practice Setting.* 2 edn. South Melbourne: The Royal Australian College of General Practitioners (RACGP), 2006.
27. Gray JA, Piribauer F. Realising the potential benefit of screening. Comment on the outcomes of the first Austrian Screening Guidelines Consensus Conference. *Wien Klin Wochenschr* 2001;113:4–6.
28. Raffle A, Gray JA. *Screening—Evidence and Practice.* New York: Oxford University Press, 2007.
29. Hofmarcher MM, Rack HM. Gesundheitssysteme im Wandel - Österreich / Health Systems in Transition - Austria. In: Riesberg A, ed, Kopenhagen: Europäischen Observatorium für Gesundheitssysteme und Gesundheitspolitik im Auftrag des WHO Regionalbüro für Europa, 2006:1–282. Health in Transition - country profiles. Figueras, J., Mossialos, E., McKee, M., and Saltman, R. B.
30. Kvas A, Kvas G, Langmann H, *et al.* *Chapter 1.2.2: "TeilnehmerInnenbefragung zur Servicequalität"*. 2008;3:6–43. VAEB Graz, Wissenschaftszentrum. Report of the central association of social insurances to the ministry of health according to §447h(4) ASVG.
31. Ladurner J, Gerger M, Holland W, *et al.* Public Health in Austria. In: London School of Economics and Political Science LH, ed. Vienna: Main Association of Austrian Social Security Institutions. Hauptverband der österr. Sozialversicherungsträger, 2010:1–262.
32. Reisner C, Dihlmann M. *Wahlärzte sichern Grundversorgung. Consilium* 10, 22–23. Wien: Niederösterreichische Ärztekammer, 2008.
33. Konsument. *Ein Blick hinter die Kulissen - Bericht Untersuchung.* 2007, 8-13. Wien: VKI, Verein für Konsumenteninformation. Tätigkeitsbericht, 2008.
34. Luciak-Donsberger C, Piribauer F, Taylor NC. Enabling general physicians to perform periodontal screening during nationwide periodic health examinations. *J Evid Based Dent Pract* 2008;2008:186–94.
35. Piribauer F. *Inhalte der neuen Vorsorgeuntersuchung—Dokumentations-Formulare (Links).* 2009. <http://www.zaeg.at/screening> (accessed 15 Mar 2009).
36. Rabe-Hesketh S, Skrondal A. *Multilevel and Longitudinal Modeling Using Stata.* 2 edn. College Station, Texas: StataCorp LP, 2008.
37. Stata Corporation. *Stata Survey Data, Reference Manual, Release 9.9th.* College Station, Texas: Stata Press, 2005.
38. Sager F. *AGITOS Zufallsgenerator.* 2010. <http://www.agitos.de/zufallsgenerator.php> (accessed 9 Jan 2010).
39. Bachinger E, Csitkovics M, Wais K. *Gesundheitsbericht Wien 2004.* Bachinger E, ed, 1994, 1-478. Wien, Stadt Wien: Gesundheitsbericht für Wien, 2005.
40. MA 5. *Tabelle 6.2.2 Niedergelassene Ärztinnen.* 2008. <http://www.magwien.gv.at/statistik/daten/gesundheitundsoziales.html> (accessed 1 Feb 2009).
41. Leyland AH, Groenewegen PP. Multilevel modelling and public health policy. *Scand J Public Health* 2003;31:267–74.
42. National Preventive and Community Medicine Committee. *Guidelines for Preventive Activities in General Practice.* 5 edn. Victoria, 2002.
43. Temml C, Piribauer F, Schmid D, *et al.* Replacing illusions with explicit targets after 30 years of a nationwide Periodic Health Examination in Austria. *Eur J Public Health* 2005;15(Suppl 1):151.
44. Deville W, van der Bij S, Groenewegen PP. *Periodic health examination by Austrian general practitioners: Developing Patient Information Forms and Health Summary Sheets.* 1 edn. Utrecht: NIVEL, 2004.
45. Paton-Simpson G, McCormick IR, Powell A, *et al.* Problem drinking profiles of patients presenting to general practitioners: analysis of Alcohol Use Disorders Identification Test (AUDIT) scores for the Auckland area. *N Z Med J* 2000;113:74–7.
46. Piribauer F, Schmid D, Klima G, *et al.* Overcoming barriers to change a 30 years lasting annual medical check up practice in Austria. *Eur J Public Health* 2005;15(Suppl 1):180. <http://www.pico.at/references/pico/eupha2005/eupha2005Posters-Contents.html>
47. Meterko M, Rubin HR. Patient judgments of hospital quality. A taxonomy. *Med Care* 1990;28:S10–14.
48. Delbanco TL. Quality of care through the patient's eyes. *BMJ* 1996;313:832–3.
49. Delbanco TL, Daley J. Through the patient's eyes: strategies toward more successful contraception. *Obstet Gynecol* 1996;88(3 Suppl):41S–7S.
50. Blendon RJ, Schoen C, DesRoches C, *et al.* Common concerns amid diverse systems: health care experiences in five countries. *Health Aff (Millwood)* 2003;22:106–21.
51. Hauptverband der österr. Sozialversicherungsträger. *Gesamtvertrag Vorsorgeuntersuchung. 343a ASVG, BGBl. Nr. 189/1955-Erlass des HVB.* 2005.
52. Hauptverband der österr. Sozialversicherungsträger. *Gesamtvertrag Vorsorgeuntersuchung inkl. Zusatzprotokoll (Gesamtvertragliche Vereinbarung).* ASVG-Erlass des HVB. 2008.
53. Tropper K. *TeilnehmerInnenzufriedenheit mit der Servicequalität bei der VU. (Satisfaction among PHE participants with their service) PDF presentation. 8-6-2010.* Vienna: VU-Träger-WS, IFGP; Institute of Health Promotion and Prevention, Austrian Social Insurance for the Railways and Mining Industry, 2012.
54. Langmann H, Tropper K, Reichenpfader U, Kvas A, Kvas G. *Aspects of Shared Decision Making and Risk Communication during the Periodic Health Examination in Austria. 8-6-2008.* Engelsberg, Schweiz: IFGP; Institute of Health Promotion and Prevention, Austrian Social Insurance for the Railways and Mining Industry, 2012.
55. McDowell I, Newell C. *Measuring Health. A Guide to Rating Scales and Questionnaires.* 1 edn. New York: Oxford University Press, 1987.
56. Palmer RH, Donabedian A, Povar GJ. *Striving for Quality in Health Care.* Ann Arbor, Michigan, 1991.
57. Grol R. Personal paper. Beliefs and evidence in changing clinical practice. *BMJ* 1997;315:418–21.
58. Deveugele M, Derese A, van dB- M, *et al.* Consultation length in general practice: cross sectional study in six European countries. *BMJ* 2002;325:472.
59. Lussier MT, Richard C. Doctor-patient communication. Time to talk. *Can Fam Physician* 2006;52:1401–2.
60. Donabedian A. Evaluating the Quality of Medical Care. *Milbank Q* 2005;83:691–729. 27-6-2012
61. Tai-Seale M, McGuire TG, Zhang W. Time allocation in primary care office visits. *Health Serv Res* 2007;42:1871–94.
62. van Dijk CE, van den Berg B, Verheij RA, *et al.* Moral hazard and supplier-induced demand: empirical evidence in general practice. *Health Econ.* Published Online First: 20 February 2012. doi:10.1002/hec.2801
63. Scott A, Connelly LB. Financial incentives and the health workforce. *Aust Health Rev* 2011;35:273–7.
64. Piribauer F. *What to do in routine primary care?—isn't it screening?* 2005. <http://www.pico.at/Hivini.htm> (accessed 17 Aug 2005).