Telephone triage by nurses in primary care out-of-hours services in Norway: an evaluation study based on written case scenarios

Elisabeth Holm Hansen,[1] Steinar Hunskaar[1,2]

▸ An additional appendix is published online only. To view this file please visit the journal online (http://qualitysafety.bmj.com).

[1]National Centre for Emergency Primary Health Care, Uni Health, Bergen, Norway
[2]Section for General Practice, Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway

Correspondence to
Elisabeth Holm Hansen, National Centre for Emergency Primary Health Care, Uni Health, Kalfarveien 31, NO-5018 Bergen, Norway;
elisabeth.holm-hansen@uni.no

Accepted 1 December 2010
Published Online First 24 January 2011

ABSTRACT

Background: The use of nurses for telephone-based triage in out-of-hours services is increasing in several countries. No investigations have been carried out in Norway into the quality of decisions made by nurses regarding our priority degree system. There are three levels: acute, urgent and non-urgent.

Methods: Nurses working in seven casualty clinics in out-of-hours districts in Norway (The Watchtowers) were all invited to participate in a study to assess priority grade on 20 written medical scenarios validated by an expert group. 83 nurses (response rate 76%) participated in the study. A one-out-of-five sample of the nurses assessed the same written cases after 3 months (n=18, response rate 90%) as a test–retest assessment.

Results: Among the acute, urgent and non-urgent scenarios, 82%, 74% and 81% were correctly classified according to national guidelines. There were significant differences in the proportion of correct classifications among the casualty clinics, but neither employment percentage nor profession or work experience affected the triage decision. The mean intraobserver variability measured by the Cohen kappa was 0.61 (CI 0.52 to 0.70), and there were significant differences in kappa with employment percentage. Casualty clinics and work experience did not affect intrarater agreement.

Conclusion: Correct classification of acute and non-urgent cases among nurses was quite high. Work experience and employment percentage did not affect triage decision. The intrarater agreement was good and about the same as in previous studies performed in other countries. Kappa increased significantly with increasing employment percentage.

BACKGROUND

Nurses' triage and telephone advice have an important place in out-of-hours services in many countries. Nurses receive calls from patients, their family or others, assess the priority grade, and decide on different actions by giving self-care advice or referring to the appropriate level of care. Several of these aspects have been reported in the literature.[1–25] Previous studies from other countries show that nurses can both underestimate and overestimate the grade of urgency.[3 10 16–18 22 23 25] Telephone triage is considered by many to be the most complex and vulnerable part of the out-of-hours services.[4 6–10]

In Norway, the municipalities are responsible for the emergency primary healthcare services, and these services include casualty clinics, primary care doctor on call and local emergency medical communication centres (LEMC). LEMCs are usually staffed with registered nurses, but in some casualty clinics the triagist may also be an enrolled nurse or a medical secretary. A registered nurse has at least a bachelor degree. An enrolled nurse or a medical secretary has 3 years in upper secondary school.

LEMCs are usually located in the casualty clinics, in the same location where the doctors are situated when they are on-call. The nurses working in casualty clinics assess the patient's condition when patients are calling and when the patient attends the clinic. In addition, the nurses are trained to assist the doctor at the casualty clinic. The nurses doing the triage on the telephone will often later meet the patients with whom they had been talking on the telephone.

The emergency medical communication centre (EMCC) is a part of the hospital level and is staffed with registered nurses. The EMCC handles the 113 calls (similar to a 999 call) and administers the prehospital emergency transportations. EMCC also alarm the LEMC and the doctor on-call when needed.

So far, no one in Norway has investigated nurses' telephone triage or the degree to which their assessments are in compliance with national guidelines.

Written case scenarios have been used in several studies to evaluate agreement with national guidelines, and they are regarded as suitable tools in the assessment of clinical competence.[8 11]

This project evaluates decisions on degree of priority made by nurses in out-of-hours services in Norway using written case scenarios. Answers were compared with consensus-based national guidelines. The intraobserver variation for the same written cases was also evaluated in a subgroup of nurses.

METHODS

The study was performed during 2008 among nurses working in seven different casualty clinics taking part in a sentinel network.

The National Centre for Emergency Primary Health Care has initiated an enterprise called 'The Watchtowers,' which is a representative sample of Norwegian municipalities and out-of-hours districts.[12] In the Watchtowers, the attending nurses record all contacts during day and night. Among other variables, they also register the degree of priority for each contact, both from telephone calls and from patients with direct attendance. All casualty clinics in the Watchtowers were invited to the study, and all agreed to participate. Altogether 116 nurses potentially could participate, but some may not have received the envelope with the study material. Eighty-eight nurses returned the material, but five were excluded from the analysis because of missing information on employment percentage, profession or number of years working in casualty clinic. Both permanent and temporary employees participated in the study. Information about initial training and instructions at the start of working in the casualty clinic were obtained from a national register.[13]

Index

The Norwegian Index for Medical Emergency Assistance is a decision tool to ensure an appropriate response to a medical emergency call.[14] The Index is available in all casualty clinics in Norway, but it is not mandatory to use it. The Index is originally intended to standardise the medical evaluation performed by nurses in EMCC. Nurses in LEMC usually know these guidelines, but they do not use them consistently for every case. This information became known by the head nurse at each WT before the study, and also told by the nurses in meetings with the researcher. Degree of priority has three designations: red colour is defined as an 'acute' response, with the highest priority; yellow colour is defined as an 'urgent' response, with a high, but lower, priority; green colour is defined as a 'non-urgent' response, with the lowest priority.[14]

Written case scenarios

Twenty written case scenarios were prepared by The National Centre for Emergency Primary Health Care (see online appendix). The distribution of degree of priority was four acute cases, eight urgent cases and eight non-urgent cases. All cases were framed in the same way regarding the patient's gender, age, time of day, mode of contact, problem/symptoms and degree of priority.

An expert group of one emergency doctor, one GP specialist and two registered nurses, one working in LEMC and one working in EMCC, evaluated the cases and compared them with the Index. The expert group classified each case with a priority grade: acute, urgent or non-urgent. There was no disagreement in the expert group about the correct response to each case, but some minor uncertainties had to be settled before the written cases were finally ready to be used.

The cases were printed in a booklet where each nurse also had to fill in profession, number of years working in the casualty clinic and employment percentage. To ensure confidentiality, each nurse entered their own code in each booklet before returning it to the Centre. Each booklet was marked in advance with a code to identify the casualty clinic.

Implementation

The leader at each casualty clinic informed the nurses about the study during staff meetings and gave each nurse an envelope containing the booklet with the written case scenarios, information about the study and a return envelope. All nurses were supposed to read and assess the cases and mark each case with the appropriate degree of priority in terms of a red, yellow or green response. The nurses were not allowed to sit together or to use the Index when assessing the cases. The answers were returned to the Centre individually.

Test–retest procedure

A sample of 20 nurses (about every fifth from the list of nurses) were to receive the same booklet after 3 months. The sequence of the cases was changed to avoid recall from the previous phase. The material was again distributed by the leader of the clinic. To identify the selected nurses, each envelope had the individual's code written outside. As an extra check of identification, the person had to fill in work-related information again (profession, employment percentage and number of years in clinic).

Statistics

SPSS version 15.0 was used to analyse data. χ^2 tests were used, and the level of statistical significance was defined

as $p<0.05$. Intraobserver variability was analysed in the sample of 18 nurses using the Cohen kappa, and agreement based on the value of κ was categorised as described by Altman.[15] The precision of κ was reported as either mean and 95% CI or median and range, and differences in κ between groups were analysed by the Kruskal–Wallis test. Over- and underestimation are weighted equally by the standard software used, as it turned out that all discrepancies were of a magnitude of 1.

RESULTS

Of the 116 potential participants, 88 returned the booklets, and the data from 83 could be analysed. The response rate thus was at least 76%. Of the 83 participants, 90% were registered nurses with or without further education, and 10% were enrolled nurses, paramedics and bioengineers. The mean duration of employment in the casualty clinic was 6.3 (median 4) years, and the mean employment percentage was 49% (median 56%). The nurses had a mean of 40 h of initial training and instruction by the start of working in the casualty clinic and LEMC, ranging from 30 to 64 h.

Assessment of case scenarios according to index

The mean of total correct responses among the nurses was 78% (SD 11.5), ranging from 45% to 95%. In total, 12% of all assessments were undertriaged, and 18% were overtriaged according to Index. Among the acute cases, 82% were correctly classified, 74% of the urgent cases were correctly classified, and 17% and 9% of the latter were overtriaged and undertriaged respectively according to Index. Of the non-urgent cases, 81% were correctly classified. One of the four acute scenarios was a patient with chest pain (case 7), and only one out of 83 nurses had undertriaged the priority grade. Table 1 shows the assessment of priority grade according to Index.

In table 2, the nurses are grouped according to the percentage of correct answers (<71%, 71–80%, >80%) and their performance is shown according to Watchtower, profession, work experience in casualty clinic and employment percentage. There were no statistically significant differences among the Watchtowers or profession. Overall, there were no significant differences between experienced or less experienced nurses in assessing the written cases, but when analysing the individual written cases, we found significant differences in three of the 20 cases (case 1, case 16 and case 17). There were no systematic differences between the experienced or the less experienced group.

Only 5% of the nurses had a full-time job, and there were no significant differences between employment percentages with respect to correct classification (table 2).

Table 1 Classification of priority grade according to index

	Acute		Urgent		Non-urgent		Total
	n	%	n	%	n	%	%
Acute							
Case 1 (n=83)	53	**64**	29	35	1	1	100
Case 7 (n=83)	82	**99**	1	1	0	0	100
Case 16 (n=82)	66	**81**	16	20	0	0	100
Case 18 (n=82)	67	**82**	15	18	0	0	100
Urgent							
Case 3 (n=82)	19	23	53	**65**	10	12	100
Case 4 (n=83)	8	9	56	**68**	19	23	100
Case 5 (n=82)	28	34	53	**65**	1	1	100
Case 9 (n=83)	10	12	66	**80**	7	8	100
Case 12 (n=83)	6	7	65	**78**	12	14	100
Case 14 (n=82)	26	32	54	**66**	2	2	100
Case 19 (n=81)	11	14	65	**80**	5	6	100
Case 20 (n=81)	6	7	70	**86**	5	6	100
Non-urgent							
Case 2 (n=83)	1	1	3	4	79	**95**	100
Case 6 (n=82)	0	0	1	1	81	**99**	100
Case 8 (n=83)	0	0	32	39	51	**61**	100
Case 10 (n=83)	0	0	4	5	79	**95**	100
Case 11 (n=83)	0	0	1	1	82	**99**	100
Case 13 (n=82)	2	2	37	45	43	**52**	100
Case 15 (n=82)	0	0	21	26	61	**74**	100
Case 17 (n=82)	4	5	22	27	56	**68**	100

Results are shown as numbers and percentages. Bold numbers indicate the percentage of correct classifications.

Table 2 Correct classifications performed by nurses according to Index in percentage groups by Watchtower, work experience in casualty clinic, profession and employment

	No of responders			
Percentage correctly classified	<71% (n=29)	71–80% (n=24)	>80% (n=30)	p Value
Watchtower				0.10
WT 1	9	2	6	
WT 2	7	5	3	
WT 3	0	2	7	
WT 4	2	6	4	
WT 5	1	1	2	
WT 6	2	4	3	
WT 7	8	4	5	
Work experience				0.30
<5 years	12	15	16	
≥5 years	17	9	14	
Profession				0.26
Registered nurses	28	20	26	
Other health professionals	1	4	4	
Employment percentage in WT				0.60
<21	6	10	6	
21–55	7	5	7	
56–75	9	4	8	
76–100	7	5	9	

Table 3 shows the percentage of correct classifications of priority grade by Watchtowers, profession, work experience and employment percentage. There was a significant difference in the proportion of correct classifications in the Watchtowers (p=0.01), but none of the other variables showed any significant differences (0.10<p<0.70). When testing the distribution of correct classifications according to priority grade, none of the independent variables in table 3 were significant (0.89<p<0.99). Differences in assessment between the casualty clinics were small, but in one single case of a pregnant woman with headache (case 4), nurses in one clinic had 100% correct assessment, while another clinic had only 47%. In the same case, there was a 20%

Table 3 Correct classifications by priority grade according to Index by Watchtower, work experience and employment percentage (results shown as percentages)

		Priority grade		
	Total	Acute	Urgent	Non-urgent
Watchtower				
WT 1	75	90	67	74
WT 2	76	73	70	82
WT 3	86	86	85	88
WT 4	80	75	81	81
WT 5	81	88	73	87
WT 6	74	77	70	77
WT 7	82	78	83	82
Work experience in casualty clinic				
<5 years	79	82	73	83
≥5 years	77	81	72	78
Profession				
Registered nurses	77	81	73	80
Other health professionals	83	86	79	85
Employment percentage				
<21	76	81	69	81
21–55	80	78	76	84
56–75	78	84	74	79
76–100	78	82	75	79
Total	78	82	74	81

overtriage of priority grade in one clinic, while another clinic had an undertriage of 41% of the same case.

Intraobserver variability

Of the subgroup of 20 nurses, 18 completed the retest of the case scenarios (response rate 90%). All were registered nurses, except one who was an enrolled nurse. The mean duration of work experience among the 18 nurses was 6.7 (median 5) years, while the mean employment percentage was 56% (median 60%).

The mean κ value for all responders was 0.61, or good (CI 0.52 to 0.70), and the range was 0.32–0.92.[15] The weighted κ value was 0.68 (good). There were no significant differences in κ values between the casualty clinics or work-experience groups, but there was a significant increase in κ value with increased employment percentage (table 4).

Owing to the small numbers, an agreement analysis within priority grades was not possible.

DISCUSSION

Strengths and limitations of the study

The use of written case scenarios cannot fully substitute for actual triage practice, but we tried to include all of the essential information needed to decide upon the priority grade—acute, urgent or non-urgent. The strength of this method consists in the fact that the nurses assess the cases on the same basis and that the situation is consistent and unchangeable. When using simulated patients or real patients, the conditions can vary because the situation is altered during the call.

Written case scenarios have been used in studies in other countries for both nurses and paediatricians working with triage.[8 16–21] An important limitation of using this methodology is that it is not assessed whether the nurses ask the patients the proper questions needed to clarify the urgency because the important information is already available. Another limitation is that the interaction skills with the patient are not challenged. When using for example simulated patients, in practice one could observe both their communication skills and their ability to ask the questions needed.[3] However, one cannot avoid differences in the presentations of the different complaints by standardised simulated patients. We think both methods (paper cases and simulation patients) have their strengths and limitations. However, in this study, we focused on the decision on priority grade given the information needed.

Correct classification of priority grade

Overall, the percentage of correct classifications in our study (78%) was higher than in most other studies. One study from The Netherlands showed correct assessments in about half of the calls,[1] while similar studies in Sweden and other countries[3 16–19] showed correct classifications from 57% to 64%. Studies from The Netherlands and USA reported about the same level of correct classification but reported a higher proportion of undertriage than we found in our study.[3 21 22]

It must be remarked that the number of priority grades used in triage varies from one country to another. Having a higher number of priority grades may lead to a lower percentage of correct classification. The fact that the nurses participating in the study both assess telephone calls and meet the patients face to face at the casualty clinic may have advantages regarding feedback of their assessment.

The underestimation of priority grades was quite low in our study. However, every error made in practice could have consequences for the patient's safety. It is therefore important that nurses obtain feedback on their priority grading to improve the accuracy.

It is commonly assumed that the Index recommends too high a priority grade in some situations (eg, sending

Table 4 Cohen kappa values, median and range by Watchtowers, duration of work experience and employment percentage

	N	Agreement	Median	Range	p Value
Watchtowers*					0.92
WT 2	5	Moderate	0.58	0.32–0.92	
WT 3	3	Moderate	0.67	0.34–0.77	
WT 4	3	Good	0.77	0.42–0.92	
WT 5	2	Good	0.62	0.47–0.77	
WT 6	5	Moderate	0.58	0.44–0.77	
Work experience					0.93
≤5 years	7	Good	0.59	0.44–0.92	
>5 years	11	Good	0.63	0.32–0.92	
Employment percentage					0.03
<34	4	Moderate	0.47	0.32–0.77	
34–70	8	Moderate	0.58	0.44–0.92	
>70	6	Good	0.77	0.67–0.92	

*Watchtowers 1 and 7 were not represented in the sample.

an ambulance when not needed). Therefore, nurses in Norway often contact the doctor on-call instead of sending an ambulance if the situation is obviously not life-threatening, thus avoiding an unnecessarily dramatic situation for the patient, and still ensuring a good and safe level of care. This may cause an undertriage according to the Index in some cases.

We found no evidence indicating that extent of work experience or nurses employment percentage affected triage decisions, which is in accordance with other studies.[16 18–21 23–27]

Intraobserver variability

There was a strong within-nurse agreement in our study. We found few studies on telephone triage that presented intrarater agreement analysed by κ. The intraobserver variability in our study was higher than that found in one study,[21] and lower than in another study[28] We had expected to find a higher level of intrarater agreement among the most experienced nurses as we found for employment percentage, but perhaps the amount of training, in this context, is more important than long work experience.

It must be remarked that both temporary and permanent employees were participating in our study. The temporary employees are working sporadically, and we showed that in the retest situation, a high employment percentage was associated with a more consistent classification of the degrees of priority. This may indicate that the temporary employees have influenced the results in a negative way.

Based on the findings from this study, it may be suggested that the LEMCs may benefit on reducing temporary employees and engage nurses in full-time jobs. Nurses in the Watchtowers both assess patients by telephone and actually meet them face to face when they attend the casualty clinic. This way of organising the casualty clinic may contribute to an important evaluation of the nurses' assessments which is of most importance regarding feedback and learning. This aspect has been given little attention in former studies. Such an attempt may strengthen the quality on decision-making and contribute to a safer service for the patients in the out-of-hour services in other countries as well as in Norway.

Telephone triage is a complex human interaction between patient and provider, and further studies are needed to assess both the quality and consistency of this activity. Advanced methodology, preferably by using real situations or experimental designs based on actors or trained patients, should be developed and validated.

Conclusions

The amount of correct classification in the three priority grades was about equal and quite high. Work experience and employment percentage did not affect triage decisions. The intrarater agreement was good and also about the same as that found in previous studies performed in other countries. From this sample of Norwegian casualty clinics, it may be suggested that the quality of decision-making is high and that nurse triage competence is safe for patients.

Acknowledgements We wish to thank all the personnel engaged in the project in the seven Watchtowers for their commitment and interest.

Funding The project is internally funded by the National Centre for Emergency Primary Health Care.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Derkx HP, Rethans JJ, Muijtjens A, *et al*. Quality of clinical aspects of call handling at Dutch out of hours centres; cross sectional national study. *BMJ* 2008;337:a1264.
2. Crouch R, Dale J. Telephone triage-identifying the demand. *Nurs Stand* 1998;12:34–5.
3. Giesen P, Ferwerda R, Tijssen R, *et al*. Safety of telephone triage in general practitioner cooperatives: do triage nurses correctly estimate urgency? *Qual Saf Health Care* 2007;16:181–4.
4. Turner VF, Bentley PJ, Hodgson SA, *et al*. Telephone triage in Western Australia. *Med J Aust* 2002;176:100–3.
5. Bunn F, Byrne G, Kendall S. The effects of telephone consultation and triage on healthcare use and patient satisfaction: a systematic review. *Br J Gen Pract* 2005;55:956–61.
6. Moll van Charante EP, ter Riet G, Drost S, *et al*. Nurse telephone triage in out-of-hours GP practice: determinants of independent advice and return consultation. *BMC Fam Pract* 2006;7:74.
7. Lattimer V, George S, Thompson F, *et al*. Safety and effectiveness of nurse telephone consultation in out-of-hours primary care; randomized controlled trial. *BMJ* 1998;317:1054–9.
8. Hasegawa T, Ogasawara C, Katz E. Measuring diagnostic competency and the analysis of factors influencing competency using written case studies. *Int J Nurs Terminol Classif* 2007;18:93–102.
9. McGrath AC, Macdonald R. Quality framework for telephone triage. *Aust Health Rev* 2008;32:786–91.
10. Car J, Koshy E, Bell D, *et al*. Telephone triage in out of hours call centres. *BMJ* 2008;337:a1167.
11. Jones TV, Gerrity MS, Earp J. Written case simulations: do they predict physicians' behavior? *J Clin Epidemiol* 1990;43:805–15.
12. Hansen EH, Hunskaar S. Development, implementation, and pilot study of a sentinel network ('The Watchtowers') for monitoring emergency primary health care activity in Norway. *BMC Health Serv Res* 2008;8:62.
13. Morken T, Hunskår S, Blinkenberg J, *et al*. Legevaktorganisering i Norge-2007. http://helse.uni.no/publications.aspx?ci=158 (accessed 18 Jan 10).
14. Norwegian Medical Association. *Norsk indeks for medisinsk nødhjelp (Norwegian Index of Emergency Medical Assistance)*. 2.1st edn. Stavanger: Åsmund S Lærdal A/S, The Laerdal Foundation for Acute Medicine, 2005.
15. Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall, 1991:403–9.
16. Bergeron S, Gouin S, Bailey B, *et al*. Comparison of triage assessments among pediatric registered nurses and pediatric emergency physicians. *Acad Emerg Med* 2002;9:1397–401.
17. Göransson K, Ehrenberg A, Marklund B, *et al*. Accuracy and concordance of nurses in emergency department triage. *Scand J Caring Sci* 2005;19:432–8.
18. Bergeron S, Gouin S, Bailey B, *et al*. Agreement among pediatric health care professionals with the pediatric Canadian triage and acuity scale guidelines. *Pediatr Emerg Care* 2004;20:514–18.
19. Maldonado T, Avner JR. Triage of the pediatric patient in the emergency department: are we all in agreement? *Pediatrics* 2004;114:356–60.
20. Worster A, Sardo A, Kevin E, *et al*. Triage tool inter-rater reliability: a comparison of live versus paper case scenarios. *J Emerg Nurs* 2007;33:319–23.

21. Belman S, Murphy J, Steiner JF, *et al*. Consistency of triage decision by call center nurses. *Ambul Pediatr* 2002;2:396–400.
22. van der Wulp I, van Baar ME, Schrijvers AJP. Reliability and validity of the Manchester Triage System in a general emergency department patient population in The Netherlands: results of a simulation study. *Emerg Med J* 2008;25:431–4.
23. Dale J, Williams S, Foster T, *et al*. Safety of telephone consultation for 'non-serious' emergency ambulance service patients. *Qual Saf Health Care* 2004;13:363–73.
24. Marklund B, Strøm M, Månsson J, *et al*. Computer-supported telephone nurse triage: an evaluation of medical quality and costs. *J Nurs Manag* 2007;15:180–7.
25. O'Cathain A, Nicholl J, Sampson F, *et al*. Do different types of nurses give different triage decisions in NHS Direct? A mixed methods study. *J Health Serv Res Policy* 2004;9:226–33.
26. Gravel J, Gouin S, Manzano S, *et al*. Interrater agreement between nurses for the pedriatic Canadian triage and acuity scale in a tertiary care center. *Acad Emerg Med* 2008;15:1262–7.
27. Olofsson P, Gellerstedt M, Carlstrøm ED. Manchester Triage in Sweden—interrater reliability and accuracy. *Int Emerg Nurs* 2008;17:143–8.
28. Rutschmann OT, Kossovsky M, Geissbuhler A, *et al*. Interactive triage simulator revealed important variability in both process and outcome of emergency triage. *J Clin Epidemiol* 2006;59:615–21.

Corrections

White RE, Trbovich PL, Easty AC, *et al*. Checking it twice: an evaluation of checklists for detecting medication errors at the bedside using a chemotherapy model. *Qual Saf Health Care* 2010;**19**:562–7.

There are two errors in the results section of this article. The authors state that "the new checklist helped nurses to detect more errors of any type (55%; 71/130) than the old checklist (38%; 49/130)". These fractions should **not** have been included because they are not a logical statistic to report. There were different numbers of planted errors in each category, making the sum of total errors unbalanced: error types which happened to have more planted errors get more weight in the fraction than those with fewer errors. The percentages reported are accurate because the authors took the average error detection percentage across each of the four types- giving them equal weight.

The authors also state that 51/60 errors in pump programming were detected with the old checklist, when it should read 54/60. The percentage value reported was correct (90%).

BMJ Qual Saf 2011;**20**:396. doi:10.1136/qshc.2009.032862corr1

Snijders C, van der Schaaf T W, Klip H, *et al*. Feasibility and reliability of PRISMA-Medical for specialty-based incident analysis. *Qual Saf Health Care* 2009;**18**:486–91.

The authors names were incorrectly cited in this paper. The author list should have been as follows; C Snijders, T W van der Schaaf, H Klip, R A van Lingen, W P F Fetter, A Molendijk.

BMJ Qual Saf 2011;**20**:396. doi:10.1136/qshc.2008.028068corr1