



Creating high-quality radiology reports in foreign languages through multilingual structured reporting

L. M. Sobez¹ · S. H. Kim¹ · M. Angstwurm² · S. Störmann² · D. Pfürringer³ · F. Schmidutz^{4,5} · D. Prezzi^{6,7} · C. Kelly-Morland^{6,7} · W. H. Sommer¹ · B. Sabel¹ · D. Nörenberg¹ · M. Berndt⁸ · F. Galiè¹

Received: 14 December 2018 / Revised: 19 February 2019 / Accepted: 26 March 2019 / Published online: 26 April 2019
© European Society of Radiology 2019

Abstract

Objectives Globalization and migration are increasing the demand for reports in different languages. We aimed to examine if structured reports created by non-German-speaking radiologists with multilingual templates show significant differences in quality to structured reports and free-text reports by German native speakers.

Methods We used structured templates that allow radiologists to report in their mother tongue and then switch the report language to German or English automatically using proprietary software. German- and English-speaking radiology residents created structured reports in both German and English with these templates. Reports for three different exam types were created (intensive care chest x-ray, shoulder x-ray specifically for degenerative processes, and CT pulmonary angiogram for pulmonary embolism). The report quality of automatically translated German structured reports by English-speaking radiologists and German structured reports by German radiologists was then evaluated by German clinicians with a standardized questionnaire. The questionnaire was designed to assess attributes including content, comprehensibility, clinical consequences, and overall quality.

Results Structured reports by English-speaking radiologists that were automatically translated into German and German structured reports by German radiologists both received very high or high overall quality ratings in the majority of cases, showing no significant differences in quality. Likewise, no significant differences were observed between the two report types regarding comprehensibility and clinical consequences. Structured reports by German radiologists received significantly better ratings for overall quality and comprehensibility compared to free-text reports by German radiologists.

Conclusions Multilingual structured reporting templates may serve as a feasible tool for creating high-quality radiology reports in foreign languages.

L. M. Sobez and S. H. Kim contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00330-019-06206-8>) contains supplementary material, which is available to authorized users.

✉ F. Galiè
Franziska.Galie@med.uni-muenchen.de

¹ Department of Radiology, University Hospital, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany

² Department of Internal Medicine IV, University Hospital, LMU Munich, Marchioninistr. 15, 81377 Munich, Germany

³ Department of Trauma Surgery, Technical University of Munich (TUM), Ismaningerstr. 22, 81675 Munich, Germany

⁴ Department of Orthopaedic Surgery, University of Munich (LMU), Marchioninistr. 15, 81377 Munich, Germany

⁵ BG Trauma Center Tübingen, University of Tübingen, Schnarrenbergstr. 95, 72076 Tübingen, Germany

⁶ School of Biomedical Engineering and Imaging Sciences, King's Health Partners, St Thomas' Hospital, King's College London, London SE1 7EH, UK

⁷ Department of Radiology, Guy's and St Thomas' NHS Foundation Trust, St Thomas' Hospital, London SE1 7EH, UK

⁸ Institute for Medical Education, University Hospital, LMU Munich, Pettenkoferstraße 8a, 80336 Munich, Germany

Key Points

- *Multilingualism in structured reporting templates can be a useful tool for creating high-quality radiology reports in foreign languages.*
- *German reports created with multilingual structured reporting templates by English-speaking radiologists and German structured reports by German radiologists exhibit no significant differences in overall report quality.*
- *Multilingual structured reporting templates can help radiologists overcome communication barriers and facilitate teleradiology.*

Keywords Radiologists · Communication barriers · Multilingualism · Teleradiology · x-rays

Abbreviations

ES	English-speaking radiologists
FTR	Free-text report
FTR_GS	Free-text reports by German speakers
GS	German-speaking radiologists
IMG	International medical graduate
RSNA	Radiological Society of North America
SR	Structured report
SR_ES	Structured reports by English speakers automatically translated into German
SR_GS	Structured reports by German speakers

Introduction

Global concerns for a shortage of healthcare workers are growing. World Health Organization projections show that by 2030, the physician shortage in countries of the Organisation for Economic Co-operation and Development (OECD) may mount to 1.2 million [1]. Many countries seek to remedy this shortage, partially by employing international medical graduates (IMGs). For instance, the number of IMGs in Germany is 11.8% and rising [2]. However, employing IMGs may also come at a cost, since language barriers and cultural differences can lead to difficulties in communication [3–7]. As such, a deficit in communication between healthcare workers has been shown to be one of the major risk factors for patient safety [8, 9]. Apart from limitations in verbal communication, language barriers could also lead to difficulties in creating and understanding written documents, such as radiology reports.

In the field of radiology, structured reporting is increasingly viewed as a tool that could potentially enable an automatic translation of reports into foreign languages [10–12]. Structured reporting templates with predefined text elements in multiple languages could allow radiologists to report in their mother tongue and then translate the report into another language automatically. While multiple studies of various exams have shown that structured reporting leads to improved completeness and comprehensibility, and results in higher satisfaction by

referring clinicians [12–23], to date, the challenge arising from reports created by non-native speakers has been addressed by only a few studies so far [24, 25].

In 2012, Stramare et al established a system for bilingual structured reporting (English/Italian) and tested this tool for the purpose of multidisciplinary meetings [24]. The authors further shared the multilingual reports on two international databases, demonstrating that multilingual structured reports (SRs) can contribute to international research and exchange of knowledge [24]. However, the generated reports had text output in bullet point form only, with an average length of 22.3 terms, and were not evaluated individually [24].

Being able to translate reports into different languages could not only facilitate teleconferences and multidisciplinary meetings as demonstrated by Stramare et al, it could also serve the growing demand for telemedicine and teleradiology. Teleradiology has the potential to improve workload distribution, on-call services, specialist consultation, shorten reporting times and reduce costs [25]. However, there is one major obstacle to international teleradiology, the language barrier. In 2010, Ross et al examined the potential of multilingual structured reporting tools to overcome this obstacle and build a cross-border teleradiology network. In this project, a multilingual structured reporting tool for knee x-rays was used to create reports in Estonian or Lithuanian which were then automatically translated into Danish [25]. The findings of these reports were then compared to those from reports created by a Danish radiologist, revealing consistent findings in 80% of reports [25].

Neither of these studies evaluated qualitative parameters of the individual reports or the satisfaction of referring clinicians. To be beneficial to patient care, reports should not only concur in their main findings, but also fulfill high-quality standards. We believe that further research evaluating the risks and potential of multilingual SRs is required to justify more widespread implementation.

We therefore aimed to compare the quality as well as the satisfaction of referring clinicians with content, comprehensibility, and clinical relevance of reports created by German-speaking and English-speaking radiologists (GS and ES) with multilingual structured reporting templates.

Materials and methods

Patient selection and study design

German-speaking and English-speaking radiologists created free-text reports (FTRs) and SRs in German and English. SRs were created and translated with multilingual structured reporting tools. Reports were evaluated with a standardized, self-designed questionnaire by clinicians.

To achieve better generalizability, we chose three different study types with varying levels of complexity: chest x-ray, shoulder x-ray, and CT pulmonary angiogram (CTPA) for pulmonary embolism.

A retrospective search was performed in our database to identify potential exams for inclusion in the present study. Image acquisition dates range from September 2013 to September 2016. Selection criteria were as follows: For supine chest x-rays, the image quality had to be adequate and at least one prior image was required. A patient's shoulder x-rays were included if they had reported shoulder pain without a history of trauma or tumor. Patients who had undergone a CTPA study that confirmed an acute pulmonary embolism were also included. From a larger pool of exams that fulfilled the eligibility criteria, 24 exams of a randomly selected subpopulation (14 men, 10 women, age 25–89 years) were included in the study.

The study was approved by our institutional review board. Informed consent was waived due to the retrospective and anonymous nature of our study.

Templates for structured reporting

Radiology reports were generated using online reporting templates which were designed using online software (Smart Reporting GmbH). The templates consist of checklists containing point-and-click menus. As the user selects different options, full sentences are generated accordingly (see Fig. 1).

Structured templates for each of the three study types were designed by experienced radiology residents under the supervision of faculty members and based on recommendations by the RSNA Reporting Initiative for CTPA [26], shoulder [27], and chest x-ray [28] (see Fig. 1 and Supplement 1). The CTPA and shoulder x-ray templates had both been evaluated in two separate, previously published studies that compared the quality of German SRs and FTRs [16, 29].

While the templates were originally created in German, they were subsequently translated into English by a medical student who is a German and English native speaker under supervision by a radiology resident. The online template included a feature that enabled the reporting radiologist to instantly switch the report language, enabling ES to create a German report (and vice versa). Before these bilingual templates were used for reporting purposes, slight adjustments

were made according to feedback by radiology residents from a British teaching hospital. While the templates contained an option to enter free text if necessary, the primary aim was to evaluate the structured templates. Therefore, readers were encouraged to limit the use of the free-text option wherever possible. In the instances where additional information was entered by the ES, the same native speaker mentioned above subsequently translated the sentences. A total of 1,848 words were translated and of these, 154 words (8.3%) were translated manually.

Creation of radiology reports

The readers were radiologists with 4 to 8 years of clinical experience including the reporting of the three exams in this study. Two were from the Department of Radiology at the University Hospital, LMU Munich (Germany) and two from the Department of Radiology at King's College London (Great Britain).

Every German reader created 8 reports for each of the three examination types, resulting in 24 reports per reader. Half of these 8 reports were SRs (SR_GS) and the other half FTRs (FTR_GS). The English readers, in contrast, only created SRs—4 reports for each of the three examination types, leading to a total of 12 reports per reader (Fig. 2a). SRs by English readers were automatically translated into German using the integrated translation tool (SR_ES).

Readers were randomly assigned to group A or B. These groups defined which 4 reports would be SRs and which 4 would be FTRs. After completing a report, the reader saved the original report and an additional, automatically translated version.

Evaluation of the reports

To avoid confounding the aptitude of the templates for creating reports in a foreign language with the language proficiency of the reporting radiologist, we evaluated the German reports, since the ES did not speak German.

The number of report evaluations for this study was based on previously published sample size calculations from prior studies that compared report quality between report types [14, 20, 23]. Following those considerations, our study sample consisted of 24 reports per report type (FTR_GS, SR_GS, SR_ES), resulting in a total of 72 reports and 144 evaluations by referring physicians.

Reports were anonymized, uniformly formatted, and inserted into a PDF document in random sequence. The age of the patient and the clinical question were included, but no clues to who had created the report.

Each report was evaluated by two German clinicians. Reports on chest x-rays ($N=24$) and CTPA for pulmonary embolism ($N=24$) were rated by two specialists for internal

Fig. 1 Screenshot of structured reporting tool. The report is generated by selecting predefined options or entering values into text fields. The translation tool allows switching the language of the template and

report flexibly [retrieved from Smart Radiology software, “CT pulmonary embolism” template, 4 Nov 2018]

medicine each (6 and 22 years of clinical experience), whereas shoulder x-ray reports ($N=24$) were rated by two specialists for orthopedics (9 and 12 years of clinical experience). The total number of evaluations amounted to 144 (Fig. 2b).

Reports were rated with an online evaluation tool called LimeSurvey [30]. Our questionnaire was created by conducting a literature review and researching the most important qualities of a radiology report [10, 21, 31–39]. We concluded that these were content, comprehensibility, and clinical consequences. These main qualities were rated on 6-point Likert scales. Additionally, we included a 6-point rating scale for overall quality and asked our clinicians to guess the nationality of the reader.

The questionnaire for report evaluation (see Fig. 3) was created in close collaboration with the LMU’s Institute for Medical Education and the physicians who later evaluated the reports. An initial draft of the online survey was discussed and modified based on their input on the wording and content of the evaluation.

Rater’s survey

After a referring physician completed the report evaluations, a follow-up survey was conducted the same day. It contained 10 free-text questions regarding the clinician’s opinion about potential advantages and disadvantages of SRs and FTRs and

which type of report the clinician would personally prefer (see Supplement 4).

Statistical analysis

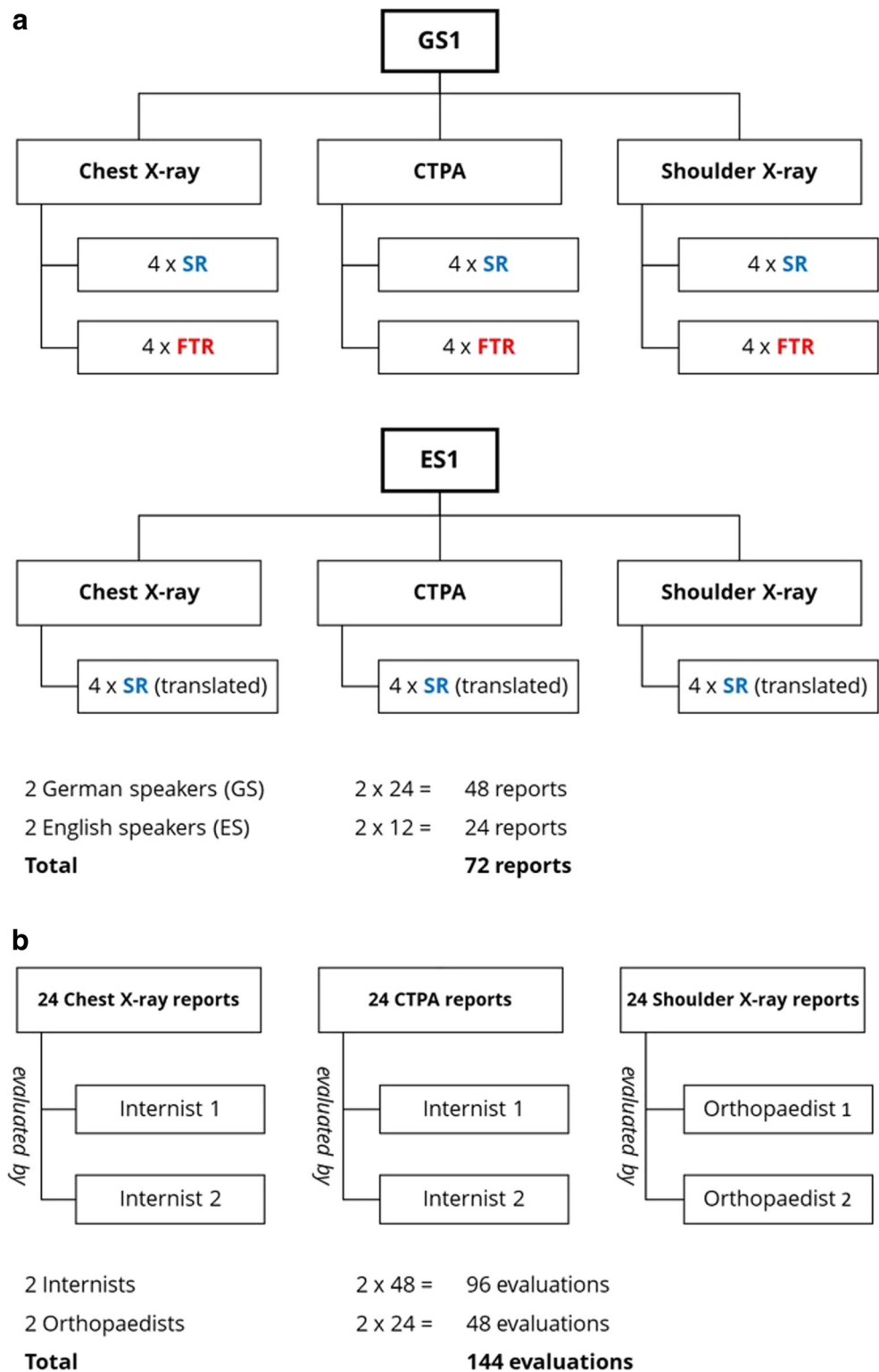
Results are summarized as medians with interquartile range or frequencies and percentages, as appropriate. Statistical analysis was performed using the non-parametric Friedman test for paired data, comparing the three different report types. Post hoc analysis with Wilcoxon signed rank tests was conducted with a Bonferroni correction applied for nine pair-wise comparison for the three items—quality, comprehensibility, and clinical consequence—resulting in a significance level set at $\alpha = 0.05/9 = 0.0056$.

Results

Overall quality

In most cases, SR_GS ($N=48$ ratings; 100%) received either very good ($N=23$; 47.9%) or good ratings ($N=11$; 22.9%). Likewise, SR_ES ($N=48$ ratings; 100%) were mostly rated to be very good ($N=15$; 31.3%) or good ($N=12$; 25.0%). In contrast, FTR_GS ($N=48$; 100%) were predominantly found to be good ($N=19$; 39.6%) or satisfactory ($N=12$; 25.0%). While among the SR_GS there was only one evaluation with a

Fig. 2 Study design. **a** Chest x-ray, shoulder x-ray, and CT pulmonary angiogram (CTPA) images were randomized and examined by two German-speaking and two English-speaking radiologists. German-speaking radiologists created 3×8 reports each (4 SRs + 4 FTRs), whereas English-speaking radiologists created only 3×4 each (4 SRs), leading to a total number of 72 reports. **b** Reports were evaluated by German clinicians. While chest x-ray and CTPA reports were reviewed by two internists, shoulder x-ray reports were assessed by two orthopaedists each



barely acceptable grade ($N = 1$; 2.1%) and unacceptable grade ($N = 1$; 2.1%) each, these numbers were higher for SR_ES (barely acceptable: $N = 5$; 10.4%, unacceptable: $N = 6$; 12.5%) and FTR_GS (barely acceptable: $N = 5$; 16.7%, unacceptable: $N = 1$; 2.1%) (see Fig. 4).

Although SR_GS received a higher number of very good or good ratings ($N = 34$; 70.8%) and a lower number of barely acceptable or unacceptable ones ($N = 2$; 4.2%) than SR_ES (very good or good: $N = 27$; 56.3%, barely acceptable or unacceptable: $N = 11$; 22.9%), an overall comparison of the

Report Evaluation

Report ID: ____

Content

The report contains enough information regarding the presence, location and severity of the following Aspects:
(1 = Strongly Disagree, 2 = Disagree, 3 = Slightly Disagree, 4 = Slightly Agree, 5 = Agree, 6 = Strongly Agree)

	1	2	3	4	5	6
Foreign material	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Congestion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pleural effusion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Infiltrate	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Pneumothorax	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Formal Aspects/ Comprehensibility

How do these Statements apply to the report?
(1 = Strongly Disagree, 2 = Disagree, 3 = Slightly Disagree, 4 = Slightly Agree, 5 = Agree, 6 = Strongly Agree)

	1	2	3	4	5	6
The report is clearly structured.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Relevant information is easy to gain.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The language of the report is clear and easy to understand. No uncommon abbreviations, vague statements or unequivocal terminology	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The report is concise.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Clinical Consequences

The report answers the clinical question.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Based on the report a decision on further clinical management (e.g. therapy, additional diagnostic tests) can be made.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Overall Report Quality

How is the overall report quality?

- Unacceptable
- Barely Acceptable
- Sufficient
- Satisfactory
- Good
- Very good

Nationality of the Radiologist

In your opinion, who created this report?

- German speaker
- English speaker
- I don't know

Fig. 3 Report evaluation questionnaire. This figure shows the report evaluation questionnaire using checkboxes and 6-point Likert scale questions for content, comprehensibility, clinical consequences, and overall report quality

report quality revealed no significant difference ($Z = -2.560$; $p = 0.010$; $\alpha = 0.0056$). In the same manner, no significant differences were found between SR_ES and FTR_GS ($Z = -1.117$; $p = 0.264$). However, SR_GS exhibited a significantly higher overall report quality than FTR_GS ($Z = -3.489$; $p < 0.001$).

A descriptive subanalysis showed differences in overall quality ratings based on the exam type. Whereas the majority of both SR_GS and SR_ES received very high or high ratings in the shoulder x-ray (87.5% of SR_GS, 93.8% of SR_ES) and CTPA exam (87.5% of SR_GS and 62.5% of SR_ES), a clearly smaller percentage of SR_GS and SR_ES were viewed to have a very high or high quality in the chest x-ray exam (37.5% of SR_GS, 12.5% of SR_ES). For chest x-ray exams, a particularly high number of SR_ES was rated as barely acceptable or unacceptable (56.3%), compared to only 6.3% of SR_GS and 6.3% of FTR_GS (see Supplement 2).

Comprehensibility

Median comprehensibility ratings added up to 5.50 for SR_GS (IQR = 1.75), 5.00 for SR_ES (IQR = 2), and 4.00 for FTR_GS (IQR = 1.75), each on a 6-point Likert scale (“The language of the report is clear and easy to understand”; 1 = strongly disagree; 6 = strongly agree). While the comprehensibility was significantly better for SR_GS compared to that for FTR_GS ($Z = -2.848$; $p = 0.004$), there were no significant differences for either of the other two combinations (see Table 1).

Clinical consequences

Median ratings for positive impact on further clinical management amounted to 6.00 for SR_GS (IQR = 2), 5.00 for SR_ES (IQR = 2), and 5.00 for FTR_GS (IQR = 2), each on a 6-point Likert scale (“Based on the report a decision on further clinical management can be made, e.g., further diagnostics or

treatment”; 1 = strongly disagree; 6 = strongly agree, see Supplement 3 for samples of reports that scored the highest possible ratings for clinical decision-making). None of the comparisons revealed a significant difference (see Table 1).

Native language of readers

For SR_GS, the evaluating clinicians were able to guess the native language of the reader correctly in 50.0% of cases ($N = 24$), while they guessed incorrectly or chose “I don’t know” in 18.8% ($N = 9$) and 31.3% ($N = 15$) of cases each.

In contrast, there were only 22.9% of correct guesses ($N = 11$) for SR_ES that were automatically translated from English into German. The referring clinicians guessed incorrectly in 41.7% of the cases ($N = 20$) and were not sure in 35.4% ($N = 17$).

Rater’s survey

All raters preferred SRs over FTRs, pointing out completeness, improved readability, and structure as the main advantages. Reduced flexibility of reports and increased length were perceived as potential drawbacks.

Discussion

Multilingual structured templates can help radiologists overcome language barriers

Our study revealed no significant differences between SR_ES (automatically translated from English to German) and SR_GS regarding overall quality, comprehensibility, and clinical consequences. Although the differences were not significant, SR_GS still received a higher percentage of very good and good reports, in addition to higher median ratings for comprehensibility and clinical consequences.

Fig. 4 Overall report quality. The majority of SR_GS and SR_ES received very good or good ratings

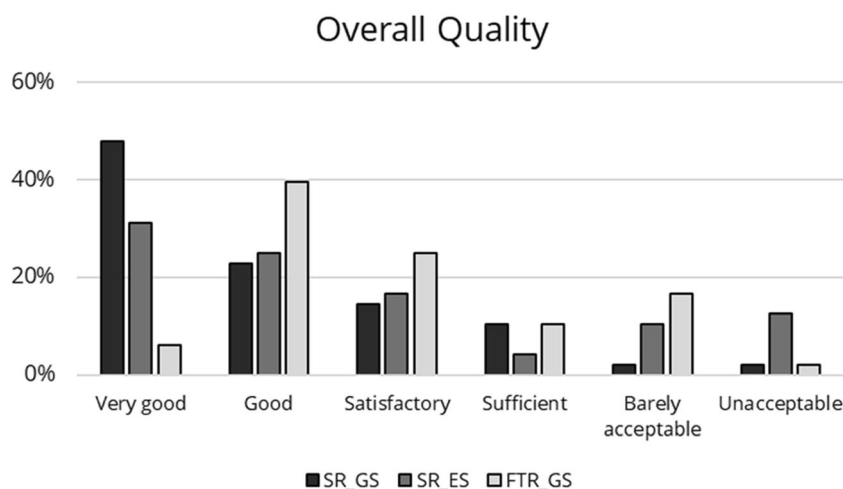


Table 1 Comprehensibility and clinical consequences. Median values and interquartile ranges of each report type for comprehensibility and clinical consequences are illustrated. SR_GS received superior medianvalues for both criteria compared to SR_ES, but the difference was not significant ($\alpha = 0.0056$). SR_GS received significantly better ratings for comprehensibility compared to FTR_GS

Comprehensibility				
Report type	Median	IQR	Min	Max
SR_GS	5.5	1.75	2	6
SR_ES	5	2	1	6
FTR_GS	4	1.75	1	6
Comprehensibility: Test Statistics ^a				
	SR_ES vs SR_GS	FTR_GS vs SR_GS	FTR_GS vs SR_ES	
Z	-1.757 ^b	-2.848 ^b	-1.474 ^b	
Asymp. Sig. (2-tailed)	0.079	0.004	0.141	
Clinical Consequences				
Report type	Median	IQR	Min	Max
SR_GS	6	2	1	6
SR_ES	5	2	1	6
FT_GS	5	2	2	6
Clinical Consequences: Test Statistics ^a				
	SR_ES vs SR_GS	FTR_GS vs SR_GS	FTR_GS vs SR_ES	
Z	-2.563 ^b	-2.045 ^b	-0.396 ^c	
Asymp. Sig. (2-tailed)	0.010	0.041	0.692	

^a Wilcoxon Signed Ranks Test^b Based on Positive ranks^c Based on negative ranks

At first glance, these results seem rather surprising: From a technical point of view, the final German report of SR_GS would be identical to that of SR_ES if only the same selection options were chosen, regardless of the input language. Additional information stated in free-text elements is also unlikely to have affected these results since there were only few cases where free text entered by the reader had to be translated manually by a native speaker.

Importantly, the differences in quality seem to predominantly concern chest x-ray reports. Nine out of 11 SR_ES with a barely acceptable or unacceptable rating were reports on chest x-ray exams. When asked what the main reason for the given grade was, reviewing clinicians most often criticized lack of pertinent information. This reason was given eight times for SR_ES and three times for SR_GS. Therefore, it seems that the ES specified fewer items, leading to less detailed reports. A possible explanation could be differences in reporting between medical centers, which a number of studies have found to be extensive [37, 40, 41]. From our personal experience, German clinicians may also expect rather detailed chest x-ray reports, whereas in surveys in other EU countries, almost 50% of clinicians thought that “no abnormal findings” could suffice as a chest x-ray report [42]. Our results on chest x-ray reports support findings by Johnson et al that in some cases, structured reporting does not necessarily lead to higher report quality [43]. Thus, it seems advisable to test each template before implementation, and the individual preferences of referring clinicians and reporting radiologists need to be taken

into consideration, especially in international research collaborations.

Irrespective of these considerations, for all three parameters, the SR_ES received at least ratings equal to those of FTR_GS, which are still used in the clinical workflow of most radiology departments [44, 45]. In addition, in most cases (77.1%), the referring physicians could not correctly guess the native language of the SR_ES, highlighting the effectiveness of the automatic translation tool. Since the utilized structured reporting software does not machine-translate free text but instead adds predefined translations for each text segment to the report, translation errors or inept wording can be avoided reliably. However, fully SR may not always be feasible (e.g., due to specific conclusions for the individual patient such as atypical findings [46]). To account for this need, combining SR with a couple of free-text sentences at the end, as a conclusion or impression, may prove beneficial for successful clinical implementation. Therefore, the templates used in this study included elements allowing for free-text entries when deemed necessary by the reporting radiologist.

Structured reporting leads to improved report quality and comprehensibility

SR_GS had a significantly higher overall quality and significantly improved comprehensibility compared to FTR_GS. All raters stated SRs as their preferred report type. These

observations are in line with findings from previous publications demonstrating the advantages of SRs compared to FTRs with regard to completeness, clarity, overall quality, and satisfaction by referring physicians [14, 16–23]. Importantly, these qualities have been confirmed in highly standardized examinations (e.g., videofluoroscopic exams [14] and cranial MRI scans for multiple sclerosis [47]) as well as in exams with a high grade of complexity (e.g., staging of rectal cancer [48] and hepatocellular carcinoma [49]).

At the same time, potential disadvantages of structured reporting need to be taken into consideration. For instance, radiologists might spend less time looking at the exam (eyeball) and more time looking at the template, which might lead to missed findings [50]. Another common concern is that the time spent creating structured reporting templates, adapting to them, and adhering to their rigid structure may lead to inefficiency [51].

Limitations

The findings of this study need to be viewed in the light of several limitations.

First, since the study had a retrospective design and reports were created in a study setting, the quality of SRs still needs to be validated in clinical routine.

Next, the assigned ratings may to some extent also reflect inter-individual differences due to the small number of readers and raters (four each). The overall quality as assessed by the raters depended both on the level of detail specified by the readers and on the personal preference of the raters. Additional prospective multi-center studies including a larger number of reporting subjects and raters may provide a more balanced view.

Also, the multilingual feature of the structured templates was only available for English and German, two closely related languages from the same language family with similar linguistic features. Therefore, the findings of this study need to be tested for additional combinations of languages, particularly ones with more extensive differences in grammar (e.g., English and Chinese).

Finally, the translation of free-text fields poses an essential challenge. Even in a SR, free-text fields are indispensable as not every potential finding can be covered by predefined elements. Although free-text fields played a minor role in the present work, the general necessity of free-text fields is a critical limitation of multilingual templates regarding their implementation in actual clinical practice. In recent years, machine-translation has emerged as a viable tool to overcome language barriers in the field of healthcare [52–54] and could potentially be used to translate free-text fields in SRs. Numerous machine-translators are already publicly available, including commercial ones (e.g., IBM Websphere Translation Server) and free, web-based ones (e.g., Bing Microsoft Translator,

Google Translate). To date, there is only very limited evidence on the translation accuracy of these applications for medical purposes. However, as a high degree of accuracy is crucial in a clinical setting, those tools warrant further evaluation before they can be used as part of the clinical workflow.

Conclusion

Altogether, these findings have important implications for the international radiological community. Due to migration and globalization, there is an increasing need for radiologists to be able to create reports in different languages in order to improve communication with international patients and colleagues. Multilingual templates could serve as an effective tool for multidisciplinary conferences, international specialist consultation, and scientific exchange in the increasingly globalized radiology community. Furthermore, the importance of telemedicine is growing and transnational teleradiology networks among countries with different languages could be facilitated, thus providing more cost-effective out-of-hour radiology services and broader coverage.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Wieland H. Sommer.

Conflict of interest The following authors of this manuscript declare relationships with Smart Reporting GmbH (online software company for structured reporting templates):

Lara Sobez: created structured templates; Su Hwan Kim: created and reviewed structured templates; Franziska Galiè: created and reviewed structured templates, background research; Wieland Sommer: co-founder.

All other authors declare no conflict of interest relevant to the study.

Statistics and biometry No complex statistical methods were necessary for this paper.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology

- retrospective
- quantitative
- single-center study

References

1. World Health Organization (2016) Global strategy on human resources for health: workforce 2030. World Health Organization,

- Geneva Available via <http://www.who.int/hrh/resources/globstrathrh-2030/en/>. Accessed 20 Sept 2018
2. Bundesärztekammer (2017) *Ärzttestistik 2017*. Bundesärztekammer, Berlin. Available via <https://www.bundesaeztekammer.de/ueber-uns/aerzttestistik/aerzttestistik-2017/>. Accessed 20 Sept 2018
 3. Slowther A, Lewando Hundt GA, Purkis J, Taylor R (2012) Experiences of non-UK-qualified doctors working within the UK regulatory framework: a qualitative study. *J R Soc Med* 105:157–165
 4. Kalra G, Bhugra DK, Shah N (2012) Identifying and addressing stresses in international medical graduates. *Acad Psychiatry* 36: 323–329
 5. Hall P, Kelly E, Dojeiji S, Byszewski A, Marks M (2004) Communication skills, cultural challenges and individual support: challenges of international medical graduates in a Canadian healthcare environment. *Med Teach* 26:120–125
 6. Skjeggstad E, Norvoll R, Sandal GM, Gulbrandsen P (2017) How do international medical graduates and colleagues perceive and deal with difficulties in everyday collaboration? A qualitative study. *Scand J Public Health* 45:428–435
 7. Klingler C, Marckmann G (2016) Difficulties experienced by migrant physicians working in German hospitals: a qualitative interview study. *Hum Resour Health* 14:57
 8. Rabol LI, Andersen ML, Ostergaard D, Bjørn B, Lilja B, Mogensen T (2011) Republished error management: descriptions of verbal communication errors between staff. An analysis of 84 root cause analysis-reports from Danish hospitals. *Postgrad Med J* 87:783–789
 9. Kohn LT, Corrigan J, Donaldson MS (2000) *To err is human: building a safer health system* (Vol. 6). Institute of Medicine (US) Committee on Quality of Health Care in America. National Academy Press, Washington, DC
 10. European Society of Radiology (ESR) (2011) Good practice for radiological reporting. Guidelines from the European Society of Radiology (ESR). *Insights Imaging* 2:93–96
 11. Bosmans JM, Neri E, Ratib O, Kahn CE Jr (2015) Structured reporting: a fusion reactor hungry for fuel. *Insights Imaging* 6: 129–132
 12. European Society of Radiology (ESR) (2018) ESR paper on structured reporting in radiology. *Insights Imaging* 9:1–7
 13. Sabel BO, Plum JL, Czihal M et al (2018) Structured reporting of CT angiography runoff examinations of the lower extremities. *Eur J Vasc Endovasc Surg* 55:679–687
 14. Schoeppe F, Sommer WH, Haack M et al (2018) Structured reports of videofluoroscopic swallowing studies have the potential to improve overall report quality compared to free text reports. *Eur Radiol* 28:308–315
 15. Buckley BW, Daly L, Allen GN, Ridge CA (2018) Recall of structured radiology reports is significantly superior to that of unstructured reports. *Br J Radiol* 91:20170670
 16. Sabel BO, Plum JL, Kneidinger N et al (2017) Structured reporting of CT examinations in acute pulmonary embolism. *J Cardiovasc Comput Tomogr* 11:188–195
 17. Norenberg D, Sommer WH, Thasler W et al (2016) Structured reporting of rectal magnetic resonance imaging in suspected primary rectal cancer: potential benefits for surgical planning and interdisciplinary communication. *Invest Radiol*. <https://doi.org/10.1097/RLI.0000000000000336>
 18. Gassenmaier S, Armbruster M, Haasters F et al (2017) Structured reporting of MRI of the shoulder - improvement of report quality? *Eur Radiol*. <https://doi.org/10.1007/s00330-017-4778-z>
 19. Brook OR, Vollmer CM, Kent TS, Sanchez N, Pedrosa I (2015) Structured reporting of multiphasic CT for pancreatic cancer: potential effect on staging and surgical planning. *Radiology* 274:464–472
 20. Sahni VA, Silveira PC, Sainani NI, Khorasani R (2015) Impact of a structured report template on the quality of MRI reports for rectal cancer staging. *AJR Am J Roentgenol* 205:584–588
 21. Schwartz LH, Panicek DM, Berk AR, Li Y, Hricak H (2011) Improving communication of diagnostic radiology findings through structured reporting. *Radiology* 260:174–181
 22. Wildman-Tobriner B, Allen BC, Bashir MR et al (2017) Structured reporting of CT enterography for inflammatory bowel disease: effect on key feature reporting, accuracy across training levels, and subjective assessment of disease by referring physicians. *Abdom Radiol (NY)*. <https://doi.org/10.1007/s00261-017-1136-1>
 23. Schoeppe F, Sommer WH, Norenberg D et al (2018) Structured reporting adds clinical value in primary CT staging of diffuse large B-cell lymphoma. *Eur Radiol* 28:3702–3709
 24. Stramare R, Scattolin G, Beltrame V et al (2012) Structured reporting using a shared indexed multilingual radiology lexicon. *Int J Comput Assist Radiol Surg* 7:621–633
 25. Ross P, Sepper R, Pohjonen H (2010) Cross-border teleradiology-experience from two international teleradiology projects. *Eur J Radiol* 73:20–25
 26. Radiological Society of North America (2014) Ct chest pulmonary embolism. Available via <http://www.radreport.org/template/0000116>. Accessed 1 Sept 2018
 27. Radiological Society of North America (2014) Shoulder xray. Available via <http://www.radreport.org/template/0000154>. Accessed 1 Sept 2018
 28. Radiological Society of North America (2014) Chest xray - post-op. Available via <http://radreport.org/template/0000145>. Accessed 1 Sept 2018
 29. Schoppe F, Sommer WH, Schmidutz F et al (2018) Structured reporting of x-rays for atraumatic shoulder pain: advantages over free text? *BMC Med Imaging* 18:20
 30. LimeSurvey GmbH (2018) The online survey tool. Available via <https://www.limesurvey.org/de/>. Accessed 5 Sept 2018
 31. Yang C, Kasales CJ, Ouyang T et al (2014) A succinct rating scale for radiology report quality. *SAGE Open Med* 2: 2050312114563101
 32. Wallis A, Edey A, Prothero D, McCoubrie P (2013) The Bristol Radiology Report Assessment Tool (BRRAT): developing a workplace-based assessment tool for radiology reporting skills. *Clin Radiol* 68:1146–1154
 33. Wallis A, McCoubrie P (2011) The radiology report—are we getting the message across? *Clin Radiol* 66:1015–1022
 34. Blackmore CC (2007) Defining quality in radiology. *J Am Coll Radiol* 4:217–223
 35. Grieve FM, Plumb AA, Khan SH (2010) Radiology reporting: a general practitioner’s perspective. *Br J Radiol* 83:17–22
 36. Kahn CE Jr, Langlotz CP, Burnside ES et al (2009) Toward best practices in radiology reporting. *Radiology* 252:852–856
 37. Naik SS, Hanbidge A, Wilson SR (2001) Radiology reports : examining radiologist and clinical preferences regarding style and content. *AJR Am J Roentgenol* 176:591–598
 38. Johnson AJ, Ying J, Swan JS, Williams LS, Applegate KE, Littenberg B (2004) Improving the quality of radiology reporting: a physician survey to define the target. *J Am Coll Radiol* 1:497–505
 39. Collard MD, Tellier J, Chowdhury AS, Lowe LH (2014) Improvement in reporting skills of radiology residents with a structured reporting curriculum. *Acad Radiol* 21:126–133
 40. Bosmans JM, Weyler JJ, Parizel PM (2009) Structure and content of radiology reports, a quantitative and qualitative study in eight medical centers. *Eur J Radiol* 72:354–358
 41. Heikkinen K, Löytyniemi M, Kormanen M (2000) Structure and content of 400 CT reports in four teaching hospitals using a new, Windows-based software tool. *Acta Radiol* 41:102–105
 42. Bosmans JM, Weyler JJ, De Schepper AM, Parizel PM (2011) The radiology report as seen by radiologists and referring clinicians:

- results of the COVER and ROVER surveys. *Radiology* 259:184–195
43. Johnson AJ, Chen MY, Swan SJ, Applegate KE, Littenberg B (2009) Cohort study of structured reporting compared with conventional dictation. *Radiology* 253:74–80
 44. Faggioni L, Coppola F, Ferrari R, Neri E, Regge D (2017) Usage of structured reporting in radiological practice: results from an Italian online survey. *Eur Radiol* 27:1934–1943
 45. Powell DK, Silberzweig JE (2015) State of structured reporting in radiology, a survey. *Acad Radiol* 22:226–233
 46. Samartine S, White L, McKeon D, Becker M (2015) Enhancing structured reporting: improving quality by tailoring the report to the clinical scenario. *J Am Coll Radiol* 12:845–847
 47. Dickerson E, Davenport MS, Syed F et al (2017) Effect of template reporting of brain MRIs for multiple sclerosis on report thoroughness and neurologist-rated quality: results of a prospective quality improvement project. *J Am Coll Radiol* 14:371–379 e371
 48. Tarulli E, Thippavong S, Jhaveri K (2015) A structured approach to reporting rectal cancer with magnetic resonance imaging. *Abdom Imaging* 40:3002–3011
 49. Flusberg M, Ganeles J, Ekinci T et al (2017) Impact of a structured report template on the quality of CT and MRI reports for hepatocellular carcinoma diagnosis. *J Am Coll Radiol* 14:1206–1211
 50. Srinivasa Babu A, Brooks ML (2015) The malpractice liability of radiology reports: minimizing the risk. *Radiographics* 35:547–554
 51. Ganeshan D, Duong PT, Probyn L et al (2018) Structured reporting in radiology. *Acad Radiol* 25:66–73
 52. Turner AM, Bergman M, Brownstein M, Cole K, Kirchhoff K (2014) A comparison of human and machine translation of health promotion materials for public health practice: time, costs, and quality. *J Public Health Manag Pract* 20:523–529
 53. Randhawa G, Ferreyra M, Ahmed R, Ezzat O, Pottie K (2013) Using machine translation in clinical practice. *Can Fam Physician* 59:382–383
 54. Dew KN, Turner AM, Choi YK, Bosold A, Kirchhoff K (2018) Development of machine translation technology for assisting health communication: a systematic review. *J Biomed Inform* 85:56–67

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.