EVIDENCE FOR DECISION MAKING IN HEALTH CARE

BASEL INSTITUTE FOR CLINICAL EPIDEMIOLOGY AND BIOSTATISTICS

Critical appraisal of the Eurofer and Swissfer study reports

Scientific evaluation report

Publication date

April 2020

Version

1.0

Authors

Dr. sc. Dominik Glinz, MSc

Prof. Dr. med. Heiner C. Bucher, MPH

Head of institute

Prof. Dr. med. Heiner C. Bucher, MPH

Table of contents

1 Aims of the critical appraisal

The objective of this critical appraisal is to systematically assess the scientific value of the Eurofer and Swissfer reports and its relevance to inform the Health Technology Assessment report on symptomatic iron-deficient patients without anemia which was commissioned by the Swiss Federal Office of Public Health (SFOPH).

The two reports to be evaluated:

1) Praxisstudie Eurofer V - Auswirkungen von individuell dosierten intravenösen Eisengaben bei Patientinnen mit Eisenmangel (Swiss Iron System SIS), SIHO Journal 2018, available on www.eurofer.ch and www.swissfer.ch (last checked 9th March 2020)
2) Effects of Individual Dosed-Intravenous Iron Doses in Patients with Iron
3) Deficiency: A Multicentre Medicine-Application Monitoring System, Schaub et al., J Gen Pract 2019, 7:2, www.swissfer.ch (last checked 9th March 2020)

The publication by Schaub et al. in the Journal of General Practice (note: this is not a recognized scientific journal, see also Section 6.2.1) is the English translation of the German report "Praxisstudie Swissfer" published on the same homepage labelled as SIHO Journal 2019. The Eurofer and Swissfer reports use the same database (Health banking) and it appears they report on the same population for almost the same observation period. As the reports report different numbers of centers and patients for similar time period (2006 to 2018 and 2006 to 2019), it seems likely that both cohorts are substantially overlapping. The evaluation in the following sections apply to both reports (Praxisstudie Eurofer V and Schaub et al. in J Gen Pract) and are called "Eurofer reports". If the evaluation resulted in an different conclusion between the two reports, this was highlighted as such (see also detailed comments to both reports in Appendix 6).

2 Methods

The evaluation of the Eurofer reports consists of a systematic assessment. Based on this systematic assessment, strengths and flaws will be addressed by bringing them into scientific and clinical context.

2.1 Systematic assessment of reporting quality

The internationally recognized initiative "to improve the reliability and value of published health research literature", the EQUATOR (Enhancing the QUAlity and Transparency Of health Research) Network (www.equator-network.org) provides guidance on how studies should be reported to ascertain the reliability and the scientific value. This includes study objective, methodology, results and further aspects. The STROBE and RECORD checklists from EQUATOR were used for the systematic assessment of the above mentioned reports. STROBE[1] provides guidance on how observational studies (Strengthening the Reporting of Observational Studies in Epidemiology) should be reported. As it is unclear whether the authors have used routinely collected health-related data, also the RECORD[2] (REporting of studies Conducted using Observational Routinely collected health Data) checklist, which is the extension of STROBE, was added.

The 22 items on the STROBE-List and the additional 13 items on the RECORD-List were evaluated point-by-point. Each point was briefly commented.

2.2 Additional items

In addition to the items on the STROBE and RECORD-list, the following items were addressed:

- Dissemination of the reports
- Legal requirements
- Data protection

2.3 Summary of the systematic evaluation of the reports

The identified strengths and major flaws will be summarized and its scientific and clinical relevance will be briefly discussed. In a last step, it will be evaluated to what extend the reports may inform the Health Technology Assessment on "Iron therapy for iron deficiency without anemia" which was commissioned by the SFOPH.[3]

3 Systematic assessment

The detailed assessment using the STROBE and RECORD checklist is presented in the Appendix 6.

4 Summary of the systematic assessment

4.1 Summary of the strengths of the reports

The authors of the reports addressed an important clinical question. They assessed the clinical effectiveness of parenteral iron therapy in women with iron deficiency from a large practices-based cohort. Unfortunately, the reporting is mainly unclear. Crucial information on data collection, population selection, choice/justification of intervention, endpoint assessment and analyses are insufficiently described. Therefore, there are no strengths that can be highlighted here beside the intention to generate real-world evidence data.

4.2 Summary of the flaws of the reports

Flaws concern:	**Explanation**
Scientific background and rationale for the investigation	Prevalences, RCT and cited references are old. At the timepoint when the two reports were «published», sufficient literature from RCTs and systematic reviews that have shown the benefit of iron therapy in females with iron deprivation were available. The authors use an own definition of iron deficiency, and ignore international and widely accepted definitions. The authors should have justified the selection of population, intervention, follow-up duration and study design: - Inclusion criteria of the selected patient population (choice of cut-off for iron indices, symptoms, symptoms severity, etc). - As the authors introduce the term "iron deficiency syndrome", the authors could also provide more background on this clinical condition. - Intervention: why was parenteral iron therapy administered instead of oral iron? - Follow-up duration: why was the follow-up limited to 3 months and not longer? - Study design: The advantages of observational studies in general practice would be: 1) to generate more generalizable results (closer

	to a real-world setting, wider patient spectrum), 2) assess longterm effect, in this particular case relapse-rates or time to relapse, 3) to assess any adverse events in a real-world setting. However, none of these points were accurately addressed. See also the following points.
Objectives	The objective of these reports is unclear as the endpoints are insufficiently defined. Moreover, the tolerability (the authors refer to "Verträglichkeit" at several occasion which they did not assess) should not be confused with investigating risk/harm.
Data collection	Data collection: The database "health banking" is insufficiently described (lacking information on funding, who entered data, data structure, data security, the purpose of the database, access, quality, validation, etc). It is unclear which data is collected (routinely collected data or were additional variables collected which are usually not collected, e.g. on symptoms). See also comment on Chapter 6.2.2 on legal requirements.
Patient selection for analysis	First: It is unclear, which patients in the Health-banking database were registered. From earlier reports, it appears that also men were registered (see for instance Eurofer 1). In later reports, it seems only women are registered. Second: It is unclear, how patient information from the health-banking database was retrieved (no selection criteria reported). It is unclear how the two reports, for almost the same period of time, reported different numbers of centers and different numbers of patients. The selection process of the patients within this database is insufficiently addressed and so, the risk of selection bias is very high. Discrepancies between two reports: Eurofer V: 107 centers (60 in CH, 36 in D, 5 in AT und 6 other countries) and 3963 patients in the periode between 2006 to 2018 Schaub et al.: 27 centers (27 in CH) and 2288 women in the period between 2006 and 2019
Outcomes	The outcome assessment, choice of outcome and presentation of results is highly problematic. Outcome assessment: 1) Who was the outcome assessor? Were outcomes patient-reported by self-administered questionnaire or by use of symptom check list by the physician or practice staff, or just by posing routine questions by the physician during consultation encounters? 2) How was the questionnaire (IDS-score) developed? How was the questionnaire structured? 3) Validation of the questionnaire? Validation in different languages? 4) It is unclear how follow-up data were collected. (follow-up visit? Phone call? Postal letter? Etc.) Choice of outcomes and definition: 5) The authors introduce their own "iron deficiency syndrome (IDS-Score)". Three symptoms of the IDS-score were not reported. Hence, the reporting of assessed and reported symptoms is inconsistent. 6) The assessed symptomes might be related (e.g. insomnia might cause fatigue, or fatigue might be the reason for concentration difficulties, etc). This has not been considered. Therefore, the

	presentation of the outcomes stratified by each symptom is problematic. 7) It is unclear whether symptom severity (continuous or categorical scale) was assessed. 8) "Success" was not predefined. From the results section, success is defined as "free of complaints or significantly improved". Other options include "slightly improved" or "unchanged". There is no option for worsening. It is unclear how "success" was assessed. Were the patients simply asked whether symptoms have improved (yes/no)? 9) How does "free of complaints, significantly improved, slightly improved or unchanged" apply to the outcome "Anemia"? Adverse events: it is unclear whether adverse events were actively registered. It is crucial that adverse events are actively assessed to make any conclusion on potential harm. From the German version "Patientinnen bekundeten Nebenwirkungen", it is not clear whether patients were actively asked, or spontaneously reported events were registered. The English version is unclear, too. Besides, it is unclear whether: - adverse events/reaction occurred during or immediately after iron infusion (the authors report only adverse events at 3 weeks follow-up), - serious adverse events were registered, - patients were excluded because of serious adverse reactions. Missing data handling is insufficiently described. Complete follow-up after 3 weeks is very suspicious and suggests that only those patients with complete data at baseline and 3 weeks were selected from a larger Health-banking database. 24% of the patients were missing at 3 months, it is unclear how this was taken into consideration in the results reported in Figure 4. Presentation of the outcomes: The graphical presentations are unclear and imprecise. - Figure 1: what do the bars show? Scale? - Figure 3: the sums of the bars do not add up to 100%, see for instant for anemia. Numbers of patients with improved symptoms should have been reported. See also above regarding unclear definition of success. - Figure 4: how was 24% of missing data taken into account? Numbers of patients with improved symptoms should have been reported. - Figure 5: Unclear how authors measured correlation. Numbers should have been reported. Unclear ferritin cut-off for 50ng/ml.
Statistical methods	Unclear: - Unclear how data was summarized. - Unclear control for confounding
Further considerations	1) The reports are disseminated via homepage and a pseudo-peer review open-access journal and hence, the publication process does not comply with scientific standard and hence the scientific rigour of the reports was probably not checked before publication (see in more detail section 6.2.1).

	2) From the reports, it is unclear whether the authors have sought ethical approval for the collection/use of patient data and whether informed consent from patients was collected. 3) It is unclear how authors ascertained data protection (e.g. anonymization of patient data, encryption of data transfers, etc).
Sections: Background and Discussion	The content in the sections Background and Discussion are written from a very narrow point of view and demonstrate the biased opinions of the study authors. The authors make claims on treatment success, low risk of parenteral iron therapy and sustainability of the therapy which are not supported by the results of their study.

5 Relevance for the HTA commissioned by the SFOPH

At first glance, the two reports seem to correspond to the population of interest in the HTA report (symptomatic, iron deficient, most are not anemic). Such a cohort might have been a great opportunity to provide complementary information on long-term effects, relapse rates or time to relapse, treatment success rates, number and type of adverse reactions caused by intravenous iron administration and number of adverse events in the real world setting (routine administration of intravenous iron). Unfortunately, the authors have missed this opportunity.

The reports have *major* limitations that affect (see sections 4.2 and 6.1 for more details):

- External validity: the patient selection is unclear, there is a high risk for bias which compromises the external validity.
- Internal validity: discrepancies of numbers of patients and centers between the Eurofer reports, and the discrepancy of inclusion criteria between older and newer Eurofer reports question the internal validity.
- Benefit assessment: the choice of outcomes does not correspond to standards in the field. The outcomes, the definition of success, outcome assessment and presentation of the outcomes is unclear. A quantification of the benefit due to parenteral iron is not possible.
- Harm assessment: it is unclear whether adverse events were actively registered. No immediate adverse reactions caused by intravenous iron administration were reported. There is a high risk of selective reporting and no quantification of potential harm due to parenteral iron is possible.

Based on the above-mentioned limitations, the Eurofer reports do not contain valuable information that would contribute valuable results to inform the clinical effectiveness or health economic assessment of the HTA report on symptomatic iron-deficient patients without anemia which was commissioned by the SFOPH.

Besides, it is highly unlikely that in future an improved reporting of the Eurofer cohort could provide scientifically sound information to inform an HTA report. Because the health-banking database structure, patient selection and outcome measures are unclear, it is questionable whether any useful information can be derived from the database.

6 Appendix

6.1 Systematic assessment using STROBE and RECORD checklists

	Item No.	STROBE & RECORD items	Praxisstudie Eurofer V, SIHO Journal 2018	Schaub et al., J Gen Pract 2019, 7:2
Title and abstract				
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Design: Yes. Summary: Partially unclear: definition of endpoints unclear. Results only reported as «beschwerdefrei», «deutlich besser», «weniger profitiert» and «unververändert». Follow-up rates? It is unclear how data was collected. Prospective / retrospective? Standardized CRFs? Regions reported. Timeframe only in the main text reported. Unclear	Design not reported with commonly used terms, probably because of poor (word-by-word) translation from German to English (unclear wording, e.g.: "progress documentation", "medicine-application monitoring"). Summary incomplete, see comments on left side. Unclear how data was registered. Hence, type of data or databases not reported in title/abstract. Geographic region reported in Abstract. Timeframe only reported in main text. Unclear
			Note: Discrepancy of reported rates for side effects: 1% vs 2.1%	

Introduction				
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Scientific content very modest. References and the prevalences described are outdated (the latest reference dates from 2003). The authors present SIHO's own definition of the severity of iron deficiency in three stages. This highly simplified classification does not allow differentiation of the different forms of iron deficiency (e.g. through chronic inflammatory reactions). The purpose of this categorization is unclear and is not further described in the reports.	See comment on left side.
			Note: When the two reports were «published», sufficient literature from RCTs and systematic reviews that have shown the benefit of iron therapy were available. **The authors should have justified the selection of population, intervention, follow-up duration and study design:** - **Inclusion criteria of the selected patient population (choice of cut-off for iron indices, symptoms, symptoms severity, etc)** - **As the Authors introduce the term "iron deficiency syndrome", the authors could also provide more background on this clinical condition.** - **Intervention: why was parenteral iron therapy administred instead of oral iron** - **Follow-up duration: why was the follow-up limited to 3 months and not longer?** - **Study design: The advantages of observational studies in general practice would be: 1) to generate more generalizable results (closer to a real world setting, wider patient spectre), 2) assess longterm effect, in this particular case relapse-rates or time to relapse, 3) to assess any adverse events in a real world setting. However, none of these points were accurately addressed. See also the following points.**	
Objectives	3	State specific objectives, including any prespecified hypotheses	The objective is unclear. Neither "typical iron deficiency symptoms" nor "success rate" were defined. It was not described how "tolerance of individually dosed	See comment on left side.

			intravenous iron treatments" was investigated, and the rest of the report did not address the tolerability.	
			Note: The objective of these reports is unclear as the endpoints are insufficiently defined. Moreover, the tolerability should not be confused with the risk/harm.	
Methods				
Study Design	4	Present key elements of study design early in the paper	The study design is described acceptable. See also point 2, the authors could have justified the study design.	Not clear: «prospective drug application monitoring»
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Setting/Locations/Rekrutierungsperioden/ Exposure/follow-up: are described, but unclear. Data collection via «Health banking» unclear: probably via the SIHO platform https://www.h-banking.com/home/ It is unclear who has access, what is recorded, and who records the data.	See comment on left side.
			Note: **Data collection: The database "health banking" is insufficiently described (lacking information on funding, who entered data, data structure, data security, purpose of the database, access, quality and validation, etc). It is unclear which data is collected (routinely collected data or were additional variable collected which are usually not collected, e.g. on symptoms). See also comment on Chapter 6.2.2 on legal requirements. In this context, it is unclear how the two reports, for almost the same period different n of centers and n of patients reported (see next line). There seems to a be a selection process of the patients within this database which is insufficiently addressed.**	
			Discrepancies between two reports: **107 centres (60 in CH, 36 in D, 5 in AT and 6 andere Länder) and 3963 patients between 2006 and 2018**	**27 centers (27 in CH) and 2288 women in the period between 2006 and 2019**
Participants	6	*Cohort study* - Give the eligibility criteria, and the sources and methods	Inclusion criteria were described, but it was not defined what are «typische	See comment on left side.

		of selection of participants. Describe methods of follow-up	Eisenmangelsymptome». It is unclear how the follow-up was done.	
		RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided.	Unclear, see above	
		RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided.	It is unclear which patients were recorded and which were considered account for the analyzes.	
		RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	It appears that no data has been linked.	
			Note: First: It is unclear, which patients in the Health-banking database were registered. From earlier reports, it seems possible that also men were registered (see for instant Eurofer 1) at the beginning. **Second: it is unclear based on which criteria information in the health-banking database were retrieved. See also point 5, the discrepancy between the n of centers and n of patients.**	

Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable. RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	The patient-relevant endpoints are insufficiently described. The purpose of the IDS Score and the categorisation into four groups is unclear. The IDS Score is no longer mentioned in the results section. While a relation between iron deficiency and most of the symptoms is possible, the clinical relation between neck tension (group 3) and iron deficiency is difficult to understand.	See comment on left side.
			Note: **Outcomes: the reporting is highly problematic:** **1) "Success" was not predefined.** **2) From the results section, success is defined as "free of complaints or significantly improved". Other options include "slightly improved" or "unchanged". There is not option for worsening.** **3) How was the questionnaire structured? Validation of the questionnaire?** **4) Outcome assessor? patient reported?** **Adverse events: only patient reported adverse events after three weeks were reported. It is unclear whether adverse events/reaction occurred during or immediately after iron infusion. It is unclear whether serious adverse events were registered.** **Exposure: In addition to the mean iron dose, the parentarel therapy could be described in more detail, especially how many injection per patient were administered.** **Predictors, confounders, and effect modifiers: were not addressed in the reports.**	
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement).	See point 7. It is unclear how the endpoints were collected.	See point 7.

		Describe comparability of assessment methods if there is more than one group		
Bias	9	Describe any efforts to address potential sources of bias	Bias was not addressed.	Bias was not addressed.
Study size	10	Explain how the study size was arrived at	The published Eurofer reports show that this is a steadily growing cohort.	See comment on left side.
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	See Point 7. It is unclear how the endpoints were summarized. The authors report frequencies and rates and in some cases state the number of patients.	
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed (e) Describe any sensitivity analyses	The description of the statistical methods is insufficient: "Die statistische Auswertung wurde gemäss Richtlinien der Biostatistik der Universität Zürich durchgeführt.» No control for confounding. Methods for subgroups and interaction are not described. There is no follow-up information for 24.2% (963/3963) of the patients after three months. It is unclear how this was taken into account when calculating success rates. No sensitivity analyzes were made.	Insufficient: «The statistical evaluation was performed pursuant to the biostatistics guidelines of the University of Zurich, Switzerland.» See comment on left side.
Data access and cleaning methods		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population.	It seems the "health banking" is handled by SIHO. However, access and cleaning methods is not described.	It seems the "health banking" is handled by SIHO. However, access and cleaning methods is not described.

		RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.		
Linkage		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Not applicable.	Not applicable.
Results				
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Surprisingly all 3963 patients were available for the 3-week follow-up (100% Follow-up). It is unclear how many women did not meet the inclusion criteria. Reasons for non-participation have not been reported. There is no flow diagram. It is unclear how the study population was selected. Were only women recruited who met the inclusion criteria?	See comment on left side.
			Note: it is unclear how many patients (men and women) were assessed for eligibility, how many are registered in the health banking database and even how many fulfilled	

			inclusion criteria for the analyses (as the reports report different numbers). See also point 7.	
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	Only age was reported. There is no follow-up information for 24.2% (963/3963) of the patients after three months. And it is uncleau how missing data was considered for analysis, e.g. figure 4. The follow-up tim point was fix.	See comment on left side.
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time	It is unclear how "success rates" was defined (see points 7 and 12). It is unclear how the endpoints were summarized. No intervals or ranges were reported. The presentation of the results in the figures and tables is unclear in many occasions. For example: Fig 1: it is unclear what the blue bars should represent. The information from Fig. 1 is also shown in Table 1. Fig. 3: the bars are inaccurate: e.g. in the case of anemia, the bars indicate more than 100% Fig. 4: unclear whether only the prevalence of the symptoms of the "successfully treated patients" were taken into account for this figure. It would have been helpful if	See comment on left side.

			the authors had also reported the number of patients. Fig. 5: Is it purely descriptive or have the authors examined the correlation? Do patients with ferritin == 50ng / ml belong to the red or light red column? Tab. 2: unclear how many injections were made to achieve ferritin T2. Text suggests that ferritin is lower after three months due to menstruation. This is pure interpretation and belongs to the Discussion section. If the authors would have had read about human iron metabolism, they should have known that after such a bolus dose, the iron first binds to ferritin and then slowly passes into other stores (e.g. liver).	
			Note: See also Point 7 and 12. Results are descriptively reported. It is unclear how the success rates in Figures 3 to 5 and reported in the text have been calculated.	
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Adjustments were not addressed.	See comment on left side.

Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	Success rates for different ferritin levels have been described. This is purely descriptive (and has not been correlated as described by the authors).	See comment on left side.
Discussion				
Key results	18	Summarise key results with reference to study objectives	Partially.	See comment on left side.
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Limitations were not addressed. Especially points 6, 7, 13 und 15 should have been addressed. It is unclear whether routine data were collected or whether the data were collected for the study-specific purpose. There is no information on the database, and hence, this point is unclear Changes to the inclusion criteria: The authors do not address this point. In older Eurofer reports (e.g. Eurofer 1), for example, men were also taken into account or a cut-off for ferritin <50 ng / ml. Changes therefore appear to have been made, but these has not been addressed. The impact and relevance of these changes is difficult to judge.	See comment on left side.
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	The interpretation oft he results goes far beyond what the results actually show. The authors neglect the discussion about possible reasons for iron deficiency,	See comment on left side.

			preventive measures, alternative therapy options, etc.	
Generalisability	21	Discuss the generalisability (external validity) of the study results	The autors do not question the external validity of their findings. This is extremely critical. The choice of the study population is unclear, especially the symptoms and the severity of the symptoms. Therefore no generalizations can be made.	See comment on left side.
Other Information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Funding is not reported.	See comment on left side.
Accessibility of protocol, raw data, and programming code		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	The authors refer to additional information on the SIHO homepage. There is no information about the protocol, raw data or programming.	See comment on left side.

6.2 Additional items

6.2.1 Dissemination of the results

The Eurofer reports I-V are available on the SIHO hosted homepage (http://www.eurofer.ch). In addition, the group has published results in the Journal of General Practice. The Journal of General Practice is open-access and articles undergo peer-review before publishing as it is described on their homepage (https://www.omicsonline.org/ArchiveJGPR/currentissue-general-practice-open-access.php) . However, the Journal of General Practice is not listed on the list of trustful open-access journals (see https://doaj.org/), nor is it indexed in Medline. Moreover, Omic has been mentioned at several occasions in relation to predatory journals (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5723186/). Predatory journals undergo no or only a pseudo-peer review, and hence do not correspond to the scientific standard of publishing. It can not be excluded whether the Journal of General Practice is a predatory journal or not, but the duration from first receipt to publication of Schaub et al. was around 5 weeks, certainly too short for proper peer-review.

Whereas the dissemination of the reports on the SIHO homepage provide a certain transparency, we think the dissemination of the results in the Journal of General Practice is highly problematic, as it pretends to be a recognized Journal and this may be misleading for lay persons.

6.2.2 Legal requirements

It is unclear whether the authors have sought ethical approval in Switzerland and the other countries.

It is unclear whether the collected data can be considered as routinely collected data or whether additional variables for the purpose of these observational cohort study were collected. In both cases ethical approval is required, however, in the latter, also informed patient consent is compulsory.

Routinely collected data can be used for research purpose without patient consent. However, authorization from the responsible ethics committees for use of health-related personal data for research purposes is required according to Article 34 “Absence of informed consent” of the Human Research Act and Art. 37-40 “Use of Biological Material and Health-Related Personal Data for Research in the Absence of Informed Consent” of the Human Research Ordinance.

Besides, it is unclear whether parenteral iron was administered according to the prescribing information (only if oral iron therapy was unsuccessful). Strictly speaking, it is unclear whether parenteral iron administration was off-label.

6.2.3 Data protection and safety

It is unclear how the database was protected (access, servers, back-up, etc), how patient privacy was assured, or whether data was encrypted or even anonymized.

7 References

1. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *Ann Intern Med.* 2007;147(8):W163-W194.
2. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med.* 2015;12(10):e1001885-e1001885.
3. Bundesamt für Gesundheit (BAG). Re-Evaluation - HTA: Iron Therapy for Iron Deficiency without Anemia - Scoping Report 2015 (14.09.2017). *https://www.bag.admin.ch/dam/bag/en/dokumente/kuv-leistungen/bezeichnung-der-leistungen/Re-Evaluation-HTA/eisentherapie-eisenmangel-ohne-anaemie-bag-scopingbericht-2015.pdf.download.pdf/Eisentherapie%20bei%20Eisenmangel%20ohne%20An%C3%A4mie%20-%20BAG-Scopingbericht%202015.pdf*. 2015.