

A STUDY SPONSORED BY THE SWISS FEDERAL OFFICE OF PUBLIC HEALTH

Longitudinal evaluation of the accuracy and completeness of clinical trial protocols – evidence for improvement?

The Adherence to SPIrit REcommendations (ASPIRE) Study

AUTHORS: DMITRY GRYAZNOV, BENJAMIN KASENDA, ERIK VON ELM, MATTHIAS BRIEL

FINAL REPORT

EXECUTIVE SUMMARY

Background: Clearly structured and comprehensive study protocols are an essential component to ensure safety and well-being of patients in clinical studies, data validity, successful study conduct, and credibility of results, particularly in the case of randomized clinical trials (RCTs). Incomplete protocols jeopardize all stages of the clinical research process with potentially harmful consequences for patients, decision-making in health care, the scientific community, and society as a whole. Funding agencies, research ethics committees (RECs), regulatory agencies, medical journals, systematic reviewers and other groups rely on protocols to appraise the conduct and reporting of the research. Evidence from cohorts of clinical trial protocols approved by RECs from the 1990s shows that RCT protocols are highly variable in their content and quality. In response to these problems, the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines were published in 2013 to improve the accuracy and completeness of clinical trial protocols through evidence-based recommendations for a minimum set of 33 major checklist items to be addressed in protocols. In January 2014 the Swiss Human Research Act and its subsidiary ordinances came into effect including a redefinition of the roles and operating procedures of Swiss RECs and Swissmedic among other points. In this context, new guidance documents for RECs and trial investigators had been developed that built on the SPIRIT framework. An evaluation of the new Swiss law has been foreseen at time of its introduction. Since improving the quality of research is one of the goals of the Human Research Act, the Swiss Federal Office of Public Health sponsored this evaluation project to investigate the accuracy and completeness of RCT protocols approved by Swiss RECs before and after the introduction of the Human Research Act (HRA).

Objectives: In this study we take the accuracy/completeness of RCT protocols as a proxy measure for the quality of RCT research in general. We focused on RCTs, because their results typically impact clinical practice and guidelines more than observational studies. Furthermore, participating patients in RCTs take on risks and burdens that impose specifically high responsibility on investigators for a considerate and professional conduct of this type of research. We defined the following specific objectives for this project:

1. To investigate the accuracy and completeness of RCT protocols approved by Swiss RECs before the introduction of the HRA in Switzerland in January 2014 (i.e. 2012) and thereafter (i.e. 2016) based on the SPIRIT checklist.
2. To evaluate the extent of registered and, in particular, prospectively registered protocols (measure of transparency) before the introduction of the HRA in January 2014 and thereafter (in national or international registries).
3. To determine trial characteristics associated with non-adherence to SPIRIT checklist items including potential interactions between year of approval (2012 vs 2016) and sponsorship of protocols (industry vs non-industry), and year of approval (2012 vs 2016) and reported methodological support from Clinical Trial Units (CTUs) or Clinical Research Organisations (CROs) (yes vs no).
4. To investigate whether accuracy and completeness of Swiss RCT protocols is different from non-Swiss RCT protocols (approved in Germany or Canada in 2012 or 2016) based on the SPIRIT checklist. In addition, we aimed to compare Swiss and non-Swiss protocols in terms of proportion of registered protocols.

Methods: For this before-after study we obtained support and established cooperation with swissethics, the seven RECs in Switzerland, and two RECs at the University Medical Centre Freiburg (Germany) and Hamilton Health Sciences (Canada). We included protocols of all approved RCTs in 2012 and 2016 in Switzerland that evaluated interventions on health outcomes. RCTs comparing different doses or routes of administration of the same drug (dose-finding studies) and trials labelled as pilot or feasibility studies were excluded. In addition, we excluded studies primarily evaluating pharmacokinetics or physiology (studies with healthy volunteers), or health economics, as well as animal studies, tissue banking, and studies using a quasi-random method of allocation. For each protocol, we first recorded information on general trial characteristics (e.g. planned sample size, single or multicentre). In a second part, we evaluated whether any information corresponding to each of the specific SPIRIT checklist items was reported (yes / no) in the RCT protocol. For protocols approved in 2012 data extraction was carried out in teams of two reviewers with methodological training working independently with subsequent agreement checks and consensus discussions in case of discrepancies. For protocols approved in 2016

we extracted only 30% of included RCT protocols independently and in duplicate for feasibility reasons. Our main outcome was adherence to SPIRIT checklist items. We calculated adherence as the proportion of trial protocols that addressed each individual SPIRIT checklist item as the mean and median number of items adhered to per protocol. We used multivariable regression analyses to investigate whether (i) year of approval (2012 vs 2016), (ii) sponsorship (industry vs non-industry), (iii) planned sample size (continuous variable), (iv) centre status (single vs multicentre), (v) reported methodological support from CTUs or CROs (yes vs no), and drug trial (vs non-drug trial) were associated with non-adherence to SPIRIT. To test our hypothesis that improvement was greater in non-industry sponsored protocols than in industry-sponsored protocols, we included the corresponding interaction term (year of approval * sponsorship) in our regression models. We used the same approach to test our hypothesis that the proportion of items adherent to SPIRIT in protocols with stated methodological support (involvement of CTU or CRO) improved less than those without. For the comparison with non-Swiss RCT protocols we included an additional independent variable (Swiss vs non-Swiss protocol) in our regression models using all available extractions from protocols approved in Switzerland, in Freiburg or Hamilton in 2012 or 2016. In addition, we carried out the interaction analyses as described above focusing exclusively on non-Swiss RCT protocols.

To assess the proportion of prospectively registered RCT protocols in 2012 and 2016 as a measure of transparency, we checked all RCT protocols and other available REC files for the documentation of a trial registration number and searched the WHO International Clinical Trials Registry Platform, clinicaltrials.gov, and Google Scholar for corresponding registration information. Then we compared the respective registration date with the corresponding enrolment date of the first participant. If the registration date was earlier or within one month of the enrolment date of the first participant, we considered the RCT as prospectively registered.

Results: We included 400 RCT protocols approved by Swiss RECs (183 RCT protocols approved in 2012 and 217 approved in 2016). Overall, we did not find a difference in the proportion of reported SPIRIT items between RCT protocols from 2012 (median, 74%, interquartile range [IQR], 64%-80%) and 2016 (median 76%, IQR, 69%-82%). However, we found a significant improvement in the subgroup of non-industry-sponsored protocols (i.e. investigator-initiated RCTs); the median proportion increased from 65% (IQR, 56%-74%) in 2012 to 76% (IQR, 66%-83%) in 2016. This improvement in non-industry RCT protocols was due to an improvement in adherence to a broad range of individual SPIRIT items and subitems with 23 individual items improving by 10% or more in the proportion of adherent protocols. In industry-sponsored protocols the proportion of reported SPIRIT items remained stable over time (median of 79%, IQR, 75%-82% in 2012 vs 77%, IQR, 72%-82% in 2016). This subgroup effect was independent of the planned size of RCTs, reported support from a CTU or CRO, intervention type (drug vs other) and single/multi-centre status. We found that the following RCT characteristics were significantly and independently associated with lower adherence to SPIRIT: single centre, no reported support from CTU or CRO, non-industry-sponsored (i.e. investigator-initiated), and approved in 2012. We did not find a subgroup effect for protocols with and without methodological support from CTUs or CROs. When we investigated 79 RCT protocols that were approved by RECs outside of Switzerland (Freiburg, Germany, and Hamilton, Canada) in 2012 and 76 RCT protocols in 2016 we found a similar subgroup effect with a modest improvement of non-industry-sponsored protocols from 2012 to 2016 (median of 59%, IQR, 53%-69% in 2012 vs 62%, IQR, 54%-70% in 2016), while the proportion of reported SPIRIT items remained about the same for industry-sponsored protocols. Although not being statistically significant there was a trend for a larger improvement of Swiss compared to non-Swiss RCTs.

We found that in both years industry-sponsored protocols were more frequently registered and more frequently prospectively registered than non-industry-sponsored protocols approved in Switzerland or outside of Switzerland (e.g. for 2012: proportion of prospectively registered protocols of 89%, 95% CI, 82%-94% for industry-sponsored Swiss protocols vs 69%, 95% CI, 58%-77% for non-industry-sponsored Swiss protocols). However, there was no evidence for an increase in the proportion of registered or prospectively registered protocols approved in Switzerland or outside of Switzerland, with one exception: The proportion of prospectively registered protocols among non-industry-sponsored

protocols approved outside of Switzerland changed from 73%, 95% CI, 56%-85% in 2012 to 83%, 95% CI, 67%-92% in 2016.

Conclusions: This before-after study suggests that the completeness of non-industry-sponsored RCT protocols approved in Switzerland improved moderately from 2012 (median 65% of SPIRIT items) to 2016 (median 76% of SPIRIT items), while industry-sponsored protocols remained on a high level without change (median 79% of SPIRIT items in 2012, and 77% in 2016, respectively). Compared with protocols approved outside of Switzerland, the improvement of non-industry-sponsored Swiss protocols appeared more pronounced, but did not reach statistical significance. There was no evidence for an improvement in the proportion of registered protocols and prospectively registered protocols in Switzerland.

This indicates an international trend for a moderate improvement in the completeness of non-industry-sponsored (investigator-initiated) RCT protocols. This is probably due to a combination of reasons, including the publication of the SPIRIT guidelines in 2013 and their implementation by research institutions, funding agencies, and medical journals; the ongoing discussion about the importance of protocol publication, thoughtful planning of RCTs, minimizing reporting biases in the scientific community; and better training in trial methodology of clinician scientists. However, we found some indications that the improvement among non-industry-sponsored RCT protocols approved in Switzerland was larger than in those approved outside of Switzerland suggesting that the HRA could have had an additional effect, most likely through guidance and templates for study protocols from swissethics, that were particularly welcomed by academic researchers. In terms of transparency more efforts are needed to enforce the prospective registration of RCT protocols in Switzerland, in particular with non-industry-sponsored protocols. This can probably be best achieved through concerted action of several Swiss stakeholders in clinical research.