

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

29.06.2018

# TABLE OF CONTENTS

LIST OF TABLES.....	3
LIST OF FIGURES.....	3
LIST OF ABBREVIATIONS.....	4
EXECUTIVE SUMMARY.....	5
1 BACKGROUND.....	8
1.1 The SPIRIT initiative.....	8
1.2 The situation in Switzerland.....	9
2 OBJECTIVES.....	9
3 METHODS.....	10
3.1 Eligibility criteria.....	10
3.2 Source of the sampled RCT protocols.....	10
3.3 Sample size considerations.....	10
3.4 Data collection.....	10
3.5 Data extraction.....	11
3.6 Confidentiality considerations.....	11
3.7 Definitions.....	11
3.7.1 Specification of variables to assess adherence to SPIRIT.....	11
3.7.2 Calculation of adherence.....	11
3.8 STATISTICAL ANALYSIS.....	12
4 RESULTS.....	14
4.1 Screening and selection of Swiss RCT protocols.....	14
4.2 Characteristics of included Swiss RCT protocols.....	14
4.3 Reporting of SPIRIT items in Swiss RCT protocols.....	16
4.3.1 Overall description of results.....	16
4.3.2 Results from multivariable regression analyses with Swiss RCT protocols.....	17
4.4 Registration of Swiss RCT protocols.....	18
4.5 Differences between Swiss and non-Swiss protocols.....	19
5 DISCUSSION.....	23
5.1 Summary of findings.....	23
5.2 Strengths and limitations.....	23
5.3 Comparison of the results with similar studies.....	24
5.4 Conclusions and future directions.....	24
6 ACKNOWLEDGMENT.....	25
7 REFERENCES.....	26
8 APPENDIX.....	27

## LIST OF TABLES

Table 1. Illustration of assigning points for adherence to SPIRIT checklist items	12
Table 2. Characteristics of all included 400 Swiss RCTs	15
Table 3. Adherence to SPIRIT items in Swiss protocols stratified by year of approval and sponsorship	16
Table 4. Results from all regression analyses based on 400 Swiss RCT protocols	17
Table 5. Registration of Swiss RCTs in trial registries	18
Table 6. Characteristics of all included 155 non-Swiss RCTs	20
Table 7. Adherence to SPIRIT items in non-Swiss protocols stratified by year of approval and sponsorship	20
Table 8. Results from regression analyses including non-Swiss RCT protocols	21
Table 9. Registration of non-Swiss RCTs in trial registries	22
Appendix Table 1. Overview of adherence to SPIRIT items of Swiss protocols using different approaches	27
Appendix Table 2. Reporting of each individual SPIRIT item stratified by year of approval and sponsorship	28
Appendix Table 3: Adherence to SPIRIT items in Swiss protocols stratified by year of approval and methodological support (yes vs no)	29
Appendix Table 4: Adherence to SPIRIT items in Swiss protocols stratified by year of approval and centre status (single vs multi-centre)	29
Appendix Table 5: Adherence to SPIRIT items in non-industry sponsored protocols stratified by year of approval and geographic area (Switzerland vs other)	29

## LIST OF FIGURES

Figure 1. Flow diagram of Swiss protocol selection for the years 2012 and 2016	14
Figure 2A. Boxplot of proportions of reported SPIRIT items by sponsorship and year	16
Figure 2B. Boxplot of proportions of reported SPIRIT items by tertiles of each year	16
Figure 3. Flow diagram of non-Swiss protocol selection for the years 2012 and 2016	19

## LIST OF ABBREVIATIONS

ASPIRE	Adherence to SPIrit REcommendations
BASEC	Business Administration System for Ethics Committees
CI	Confidence Interval
CRO	Clinical Research Organisation
CTU	Clinical Trial Unit
EKNZ	Ethics Committee of Northwestern and Central Switzerland
EudraCT	European Clinical Trials Database
FOPH	(Swiss) Federal Office of Public Health
HRA	(Swiss) Human Research Act
HTA	Health Technology Assessment
ICMJE	International Committee of Medical Journal Editors
IQR	Interquartile range (25 <sup>th</sup> -75 <sup>th</sup> percentile)
KEK-ZH	Ethics Committee of Zurich
NA	Not Applicable
NIHR	National Institute for Health Research
OR	Odds Ratio
RCT	Randomised clinical trial
REC	Research ethics committee
SAMS	Swiss Academy of Medical Sciences
SCTO	Swiss Clinical Trial Organisation
SNSF	Swiss National Science Foundation
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials
Swissmedic	Swiss Agency for Therapeutic Products
UK	United Kingdom

## 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.

# 1 BACKGROUND

Every randomised clinical trial (RCT) has to be based on a protocol, a document that details the study rationale, proposed methods, organisation and ethical considerations (1). The protocol serves as the foundation for planning, conduct, reporting, and appraisal of the study. If the clinical trial protocol is inadequate or lacks essential information about how the trial will be conducted, this can have a fundamental impact on the future trial results. 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 (2). To meet the needs of all stakeholders (including the trial investigators themselves), protocols should adequately address the key elements of clinical trials. At present, RCT protocols are highly variable in their content and quality (3). Previous empirical studies assessing the content and quality of RCT protocols show that many do not adequately describe the primary outcomes (inadequate for 25% of trials), treatment allocation methods (54-79%), use of blinding (9-34%), adverse event reporting methods (41%), components of sample size calculations (4-40%), data analysis plans (20-95%), publication policies (7%), or the roles of sponsors and investigators in study design or data access (89-100%) (4–9). These studies are based on evidence from cohorts of clinical trial protocols approved by RECs from the 1990s. As such, we have no up-to-date evidence on the current accuracy and completeness of RCT protocols.

An inadequate and incomplete RCT protocol has major implications for the clinical trial team, patients involved in the trial, funders and reviewers. Most importantly it can lead to increased (compared to the same study based on an adequate protocol) risks for participants as well as to useless and misleading results being applied to the detriment of patients, and last but not least in huge amounts of money invested in research being wasted (10). Lack of detail in the key elements of the protocol can adversely affect the conduct of the trial. For example, if the outcomes to be measured are not precisely defined, it is difficult for study personnel to adequately collect outcome data. Missing or unreliable data threaten the validity of trial results and risk wasting the contribution made by patients and the trials team as well as the investment made by funders and sponsors (11). We are not aware of any direct evidence linking poor quality of study protocols to poor patient outcomes, but extrapolation of existing evidence and scientific logic strongly suggest this relationship. Furthermore, deficiencies in the original protocol can lead to avoidable protocol changes and associated delays. It is estimated that about half of protocols approved by French RECs have subsequent amendments (12). A third of amendments submitted for industry-sponsored trial protocols initiated in 2006 to 2008 were classified by the sponsor as avoidable with greater attention to the protocol (13). Each amendment introduces study delays of between 6 to 16 weeks on average entailing additional costs and burden for the REC.

Incomplete protocols also negatively affect the process of external review. When key elements of trial methodology are missing from the protocol, it is unclear to the reviewer whether they were not adequately addressed by the clinical trial team, or adequately addressed but not documented in the protocol. This lack of clarity creates potential delays in study approval as investigators are contacted for further clarification. For example, French RECs requested revisions to about half of submitted protocols before approval. After trial completion, journal editors, peer reviewers, and systematic reviewers will be unable to adequately appraise the trial and identify any discrepancies with the protocol e.g. if it is unclear whether study components such as a subgroup analysis were pre-specified or not (2,14).

In summary, clearly structured and comprehensive study protocols are essential to ensure the safety and well-being of trial participants, data validity, successful study conduct, and credibility of results, particularly in the case of RCTs. Incomplete protocols jeopardize all stages of the clinical research process with harmful consequences for patients, decision-making in health care, the scientific community, and society as a whole.

## 1.1 *The SPIRIT initiative*

In response to these problems, the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) Initiative was launched in 2007. This international collaborative project aims to improve the completeness of clinical trial protocols by producing evidence-based recommendations for a minimum set of checklist items to be addressed in protocols. The SPIRIT recommendations were developed using systematic, transparent methodology and broad consultation with 115 experts representing diverse stakeholders involved in the design, funding, conduct, review, and publication of trial protocols. The SPIRIT Statement with its 33 items and the accompanying Exploration and Elaboration paper were published in January 2013 (11,15). SPIRIT builds on other applicable international guidance by citing empirical evidence and providing additional recommendations. It adheres to the ethical principles mandated by the 2008 Declaration of Helsinki (1), and encompasses the protocol items recommended by the International Conference on Harmonisation Good Clinical Practice E6 guidance. Since its publication, SPIRIT has been widely endorsed by key stakeholders, including over 75 journals (e.g., Lancet, BMJ, Annals of Internal Medicine), the World Association of Medical Editors, drug manufacturers (e.g. GlaxoSmithKline, Johnson and Johnson), and research institutions (e.g. MRC



Clinical Trials Unit, and the NCIC Clinical Trials Group) ([www.spirit-statement.org/about-spirit/spirit-endorsement/](http://www.spirit-statement.org/about-spirit/spirit-endorsement/)). An essential component of the implementation of SPIRIT is to evaluate the impact of the guideline on the quality of protocol reports over time.

## **1.2 The situation in Switzerland**

Stakeholders of clinical research in Switzerland such as the Federal Office of Public Health (FOPH), the Swiss National Science Foundation (SNSF), the Swiss Academy of Medical Sciences (SAMS), and the University Hospitals have recognized the need for infrastructure and know-how to facilitate the conduct and improve the quality of clinical research. As one consequence, the SNSF and the Swiss Academy of Medical Sciences created six clinical trial units and the Swiss Clinical Trial Organisation (SCTO) more than 10 years ago. In January 2014, the new federal Law on Research in Humans (Human Research Act, HRA) and its subsidiary ordinances came into effect, implying a redefinition of the roles and operating procedures of the Swiss RECs and the drug licensing authority Swissmedic. In this context, new guidance documents for ethics committees and trialists had been developed that are built on the SPIRIT framework ([www.swissethics.ch](http://www.swissethics.ch)). The new Swiss law foresees an evaluation of itself and its introduction; and since quality of research is one of the goals of the HRA, the Swiss FOPH sponsored an evaluation project in which we designed a retrospective cohort study on RCT protocols approved by Swiss and non-Swiss RECs.

## **2 OBJECTIVES**

In the present 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 and participants in RCTs take on risks and burdens that impose high responsibility on trial investigators for a considerate and professional conduct of this type of research. We defined the following specific objectives for the present study:

1. To investigate the accuracy and completeness of RCT protocols approved by Swiss RECs two years before (2012) the introduction of the HRA in Switzerland (January 2014) and two years thereafter (2016) based on the SPIRIT checklist.
2. To evaluate the extent of registered and, in particular, prospectively registered protocols (measure of transparency) before (in 2012) the introduction of the HRA and thereafter (in 2016) 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 or 2016) and sponsorship of protocols (hypothesis: Investigator-sponsored protocols improved while industry-sponsored protocols did not change), and year of approval (2012 or 2016), and reported methodological support from CTUs or Clinical Research Organisations (CROs) (hypothesis: methodologically supported protocols improved less than RCT protocols without reported methodological support).
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; i.e. in case we find a difference in completeness between Swiss protocols from 2012 and 2016, we wanted to know, whether this difference is unique for Swiss protocols or whether there is evidence for an international trend. In addition, we aimed to compare Swiss and non-Swiss protocols in terms of proportion of prospectively registered protocols.

### 3 METHODS

To meet our objectives, we established research collaborations with key people of the SPIRIT initiative (Prof Doug Altman & Dr. Sally Hopewell, Centre for Statistics in Medicine at the University of Oxford, UK, and Prof An-Wen Chan, Women's College Research Institute at the University of Toronto, Canada). Briefly, we conducted the following steps:

- First, we established a cohort of RCT protocols approved in 2012 prior to the publication of the SPIRIT guideline and the enactment of the Swiss HRA in January 2014.
- Second, we established a cohort of RCT protocols approved in 2016 (i.e. after publication of SPIRIT and after the enactment of the Swiss HRA)
- Third, we compared the two cohorts regarding reporting and adherence to the defined 33 major SPIRIT items with subitems to assess potential improvement in RCT protocol completeness over time.

#### **3.1 Eligibility criteria**

We included protocols of all approved RCTs in 2012 and 2016 in Switzerland that compared an intervention with placebo, a sham intervention, another active intervention or no intervention or combinations thereof. We defined an RCT as a prospective study in which patients, or groups of patients, were assigned at random to one or more interventions to evaluate their effects on health outcomes. Studies comparing different doses or routes of administration of the same drug (dose-finding studies) and trials labelled as pilot or feasibility studies were included but represent two pre-specified subgroups, which are not included in the analyses reported herein. We also excluded studies primarily evaluating pharmacokinetics or physiology (studies with healthy volunteers), or health economics, as well as animal studies, tissue banking, observational studies, studies involving only qualitative methods, and studies using a quasi-random method of allocation (such as alternation, date of birth, or case record number).

#### **3.2 Source of the sampled RCT protocols**

We obtained support and established cooperation with swissethics and the seven RECs in Switzerland as well as RECs at the University Medical Centre Freiburg, Germany, and the Hamilton Health Sciences, Ontario, Canada, building on successfully completed previous research (16).

#### **3.3 Sample size considerations**

Based on the above eligibility criteria and meta-data from Swiss RECs from 2012, we estimated that there would be about 250 eligible RCT protocols from Swiss RECs for each year (2012 and 2016) to be included in our analyses (total of 500 RCT protocols approved by Swiss RECs). For feasibility reasons when there were large numbers of eligible RCT protocols (more than 70 protocols in one year) from specific RECs, we decided to draw a stratified random sample of 60 RCT protocols from these specific RECs. The stratification was by tertile of the submission date in 2012 and 2016 (20 protocols from each time period: Jan-April; May-Aug; Sept-Dec). For the present study, we drew stratified random samples of 60 RCT protocols for the REC in Zurich for 2012 and the RECs in Northwestern Switzerland and Zurich for 2016. In addition to the Swiss RCT protocols we drew random samples of 45 eligible RCT protocols from the RECs in Freiburg (Germany) and Hamilton (Canada) for each year (2012 and 2016 with 15 protocols for each time period: Jan-April; May-Aug; Sept-Dec; total of 90 RCT protocols for both years). In principle, there are further RCT protocols from Southern England that could be included in a pooled analysis in the future, however, protocols from Southern England have only been assessed from 2012 so far. For the present analysis, we focus on RCT protocols approved by one of the seven Swiss RECs and a comparison with a random sample of 45 RCT protocols each from Freiburg (Germany) and Hamilton (Canada) from 2012 and from 2016 (total of 155 eligible non-Swiss protocols).

#### **3.4 Data collection**

For each protocol, we first recorded information on general trial characteristics including the investigator names, sponsor, funding sources, medical specialty field, type of interventions, type of patients, number of study centres, number of study arms, study design and planned sample size. In a second part, we evaluated whether any information corresponding to each of the specific individual SPIRIT checklist [15] items was reported (yes / no) in the RCT protocol. The general trial characteristics and specific questions about SPIRIT items and sub-items were discussed and pilot-tested by two researchers from Basel (Briel M, von Niederhäusern B) and two researchers from the UK (Oduyayo A, Hopewell S) until we reached consensus about a core set of variables and their definitions.

### 3.5 Data extraction

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. To ensure consistency in the data extraction process, all reviewers first received training on how to extract each data point. They then extracted data from a sample of 3-5 RCT protocols which were checked for agreement with data extracted by a member of the core team (Gryaznov D, von Niederhäusern B, Briel M); any discrepancies in the data were discussed with each reviewer (calibration process). Following this calibration process, we extracted data independently and in duplicate for 95% of included RCT protocols from 2012 and 30% of included RCT protocols from 2016. The remaining 5% of protocols from 2012 and 70% of protocols from 2016 were extracted by one reviewer only. This difference in duplication rate is due to limited resources only. We used a web-based password-protected data extraction tool (<http://www.squiekero.org>) for data entry, storage, and data management that we used in prior research (14,16).

### 3.6 Confidentiality considerations

The Swiss, German, and Canadian RECs were all project partners and we collaborated in the same way as we did before (16) with a mandate from each participating REC for 2012 and from swissethics for Swiss protocols from 2016. All researchers extracting data from RCT protocols signed confidentiality agreements with RECs or Swissethics to conduct the outlined project within quality assurance measures of Swissethics and to confidentially handle the information contained in REC files which did not leave the REC offices. Only aggregated data are to be published and none of the primary studies, investigators, or sponsors are identifiable. The final database contained only data with coded trial identification numbers.

### 3.7 Definitions

#### 3.7.1 Specification of variables to assess adherence to SPIRIT

The SPIRIT checklist includes 33 different major items. To measure adherence in trial protocols, we pre-specified a total of 64 variables to be extracted from each trial protocol, because some SPIRIT items have several components and some explicit sub-items (e.g. #5a-d). These 64 variables have the values: “Yes”, “No”, or (sometimes) “Not applicable”. Different scenarios in terms of data structure are possible depending on the complexity of each SPIRIT item:

- **Single SPIRIT items (Type 1 variables).** SPIRIT items for which we extracted a single variable (n=19 items and n=19 variables in total; SPIRIT items number 1-4,7-9, 13, 22-25, 27-30, 32-33, 19).
- **Multiple component items (Type 2 variables).** SPIRIT items for which we extracted more than one variable (n=4 SPIRIT items: 10 (3 variables), 12 (3 variables), 14 (6 variables), 15 (3 variables)).
- **Multiple explicit sub-items (Type 3 variables).** SPIRIT items which consist of multiple sub-items and for which we extracted one variable for each sub-item (n=8 SPIRIT items: 5 (a-d), 6 (a, b), 16 (a-c), 18 (a, b), 20 (a-c), 21 (a, b), 26 (a, b), 31 (a-c)).
- **Multiple explicit sub-items with several components in sub-items (Type 4 variables).** SPIRIT items which consist of multiple sub-items and for which we extracted more than one variable for one of the sub-items (n=2 SPIRIT items: 11 (a-d) with 2 variables for 11a, and 17 (a, b) with 3 variables for 17a).

#### 3.7.2 Calculation of adherence

To calculate adherence to the SPIRIT checklist, we used three different approaches:

1. **Major item approach (simple):** For each of the 33 major SPIRIT items we assigned one point for each “Yes” or “Not applicable” in Type 1 variables, and one point **if all** Type 2, Type 3 and Type 4 variables contingent to a major SPIRIT item were “Yes” or “Not applicable”. Otherwise zero points were assigned. The maximum possible score with this approach was 33 points.
2. **Major item approach (allowing for partial credit of individually met subitems or components of major SPIRIT items):** For each Type 1 variable we assigned one point for each “Yes” or “Not applicable”. We assigned a fraction of one point for each sub-variable in Type 2 and Type 3 variables. For example, if there were two sub-variables, each got 0.5 points for a “Yes”. In case there were 3 sub-variables, each got 1/3 point. For Type 4 variables we applied the same rule, i.e. for example item #17 consists of 17a and 17b. Each of these were assigned 0.5 points in case of a “Yes” or “Not applicable”; however, #17a consists of three components, and therefore each of these type 4 variables were assigned 1/3 of 0.5 (=0.1667) points in case of a “Yes” or “Not applicable”. A “No” led to zero points in each case. The maximum possible score with this approach was 33 points.

3. **All item approach:** For each “Yes” or “Not applicable” in each variable (Type 1, 2, 3, or 4) we assigned one point. A “No” was assigned zero points. The maximum possible score with this approach was 64 points.

In a sensitivity analysis we repeated the calculations with each of the mentioned approaches but assigned points only in case of a “Yes” for a specific variable; in case of “Not applicable” we assigned neither one nor zero points, but did just not consider this item for a specific protocol. This means that the maximum possible score could vary across protocols for each of the three approaches.

Item	Value	Assigned points according to approach		
		Major items (simple)	Major items (allowing for partial credit)	All items
1	Yes	1	0.1667	1
1a.1	No	-	0	0
1a.2	Yes	-	0.1667	1
1a.3	No	-	0	0
1b.1	Yes	-	0.5	1
2	Yes	0	0.5	1
2a	No	-	0	0
2b	Yes	-	0.5	1
3	No	0	0	0
4	Yes	1	1	1
4a.1	Yes	-	0.1667	1
4a.2	Yes	-	0.1667	1
4a.3	Yes	-	0.1667	1
4b.1	Yes	-	0.1667	1
4b.2	Yes	-	0.1667	1
4b.3	Yes	-	0.1667	1

**Table 1:** Illustration of assigning points for adherence to SPIRIT checklist items according to different approaches

### 3.8 STATISTICAL ANALYSIS

Our main outcome was adherence to SPIRIT checklist items reported in RCT protocols approved by Swiss RECs in 2012 and in 2016. We calculated adherence as the proportion of trial protocols that addressed each individual SPIRIT checklist item (according to the different approaches described above) as the mean / median number of items adhered to per protocol. Our main analyses was based on the second approach (major item approach allowing for partial credit of individually met subitems or components of major SPIRIT items) with “not applicable” getting assigned a point, because it keeps the hierarchical structure of the SPIRIT checklist and it independently considers all components and sub-items of all individual SPIRIT items. For descriptive analyses and data presentations in tables, we stratified included protocols by the year of approval (2012 versus 2016) as well as by sponsorship (industry versus investigator). To investigate whether the following variables (independent variables) were associated with a larger proportion of SPIRIT items adhered to (dependent variable), we used multivariable regression models (specific objectives 1 & 3):

- Year 2012 *versus* 2016 (Hypothesis: RCT protocols approved in 2016 are more comprehensive due to available protocol templates from swissethics introduced together with the HRA in 2014, thus associated with a larger proportion of adherence)
- Investigator sponsorship *versus* industry sponsorship (Hypothesis: industry-sponsored RCT protocols are more comprehensive, thus associated with a larger proportion of adherence)
- Sample size (in steps of 100) (Hypothesis: larger RCTs are better planned with more comprehensive protocols, thus associated with a larger proportion of adherence)
- Single centre *versus* multi-centre RCTs (Hypothesis: multi-centre RCTs are better planned with more comprehensive protocols, thus associated with a larger proportion of adherence)

- Lack of methodological support *versus* support from CRO or CTU (Hypothesis: protocols mentioning the support of a CRO or CTU are more comprehensive, thus associated with a larger proportion of adherence)
- Drug *versus* non-drug intervention trials (Hypothesis: drug trials are better planned due to the more stringent regulatory requirements and, therefore, have more comprehensive protocols associated with a larger proportion of adherence)

Since different types of regression analyses all have their specific advantages and disadvantages, in the analysis of the present study we used linear, beta (17), and hierarchical logistic regression models. In the linear and beta regression models we used the proportion of the SPIRIT items adhered to per protocol as the dependent variable. In the hierarchical logistic regression model we considered two data levels: the “SPIRIT item level” and the “protocol level”. Each SPIRIT item could assume the values zero or one (logistic dependent variable) and all items were clustered by protocols. Therefore, we included independent variables in the hierarchical logistic model as fixed effects and protocols as a random effect. For all types of regression analyses we reported linear coefficients or odds ratios (ORs) accompanied by 95% confidence intervals (CIs).

To specifically test our hypotheses that investigator-sponsored protocols improved in terms of adherence to SPIRIT between 2012 and 2016 while industry-sponsored protocols did not (because of less room for improvement with industry-sponsored protocols), we included a corresponding interaction term (year of approval \* sponsorship) in each of the mentioned multivariable regression models. We used the same approach to test our hypothesis that methodologically supported protocols (involvement of CTU or CRO) improved less than RCT protocols without reported methodological support.

To assess the proportion of registered RCT protocols in 2012 and 2016 we carefully checked all RCT protocols and other available REC files for the documentation of a trial registration number. If we could not find a registration number with the RCT protocol or associated files we searched the WHO International Clinical Trials Registry Platform ([www.who.int/ictpr](http://www.who.int/ictpr)) for a corresponding registration number. In case we did not find an RCT protocol on the WHO registry platform we searched [clinicaltrials.gov](http://clinicaltrials.gov), the European Clinical Trials Database (EudraCT), and google scholar, and only if we did not find any valid registration number then, we categorized the RCT protocol as “not registered”. Of the registered protocols we extracted the date of first registration and the date of entry of the first RCT participant from the registry. We defined “prospectively registered protocols” as protocols with a date of first registration within a month of the entry date of the first RCT participant to allow for processing delays in the registry (18). We calculated proportions of “RCT protocols registered at all” and “prospectively registered RCT protocols” for 2012 and 2016 accompanied by 95% CIs.

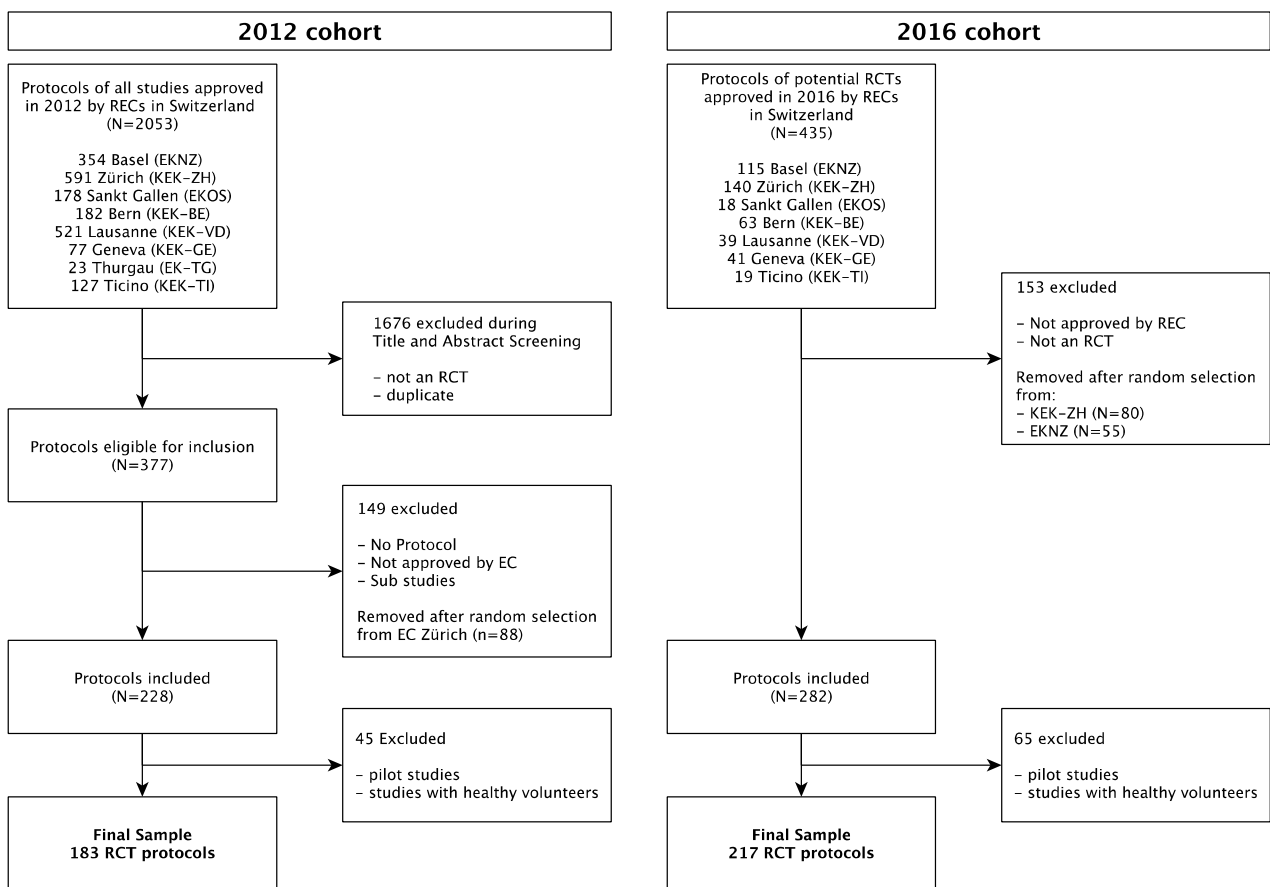
For the comparison with non-Swiss RCT protocols (objective 4) we included an additional independent variable (Swiss vs non-Swiss protocol) in our regression model using all available extractions from protocols approved in Switzerland or in Freiburg (Germany) or Hamilton (Canada) in 2012 or 2016. In addition, we carried out the interaction analyses as described above focusing exclusively on non-Swiss RCT protocols. All data analyses were performed at the significance level of 0.05 if not specified otherwise.

## 4 RESULTS

This report summarizes the findings for specific Objectives 1 to 4 focusing exclusively on RCT protocols approved by Swiss RECs or RECs in Freiburg (Germany) and Hamilton (Canada). Results from RCT protocols in the UK will follow but are outside the scope of this report.

### 4.1 Screening and selection of Swiss RCT protocols

We have screened 2488 trial protocols (2053 from year 2012 and 435 from year 2016) at archives of all seven Swiss RECs (in 2012 there were eight Swiss RECs, but St. Gallen and Thurgau were merged in the subsequent years for a total of seven RECs). After exclusion of ineligible protocols, we finally included 400 approved RCT protocols (183 from year 2012 and 217 from year 2016) for our analyses. Details of the selection process are provided in **Figure 1**.



**Figure 1:** Flow of RCT protocol selection for the years 2012 and 2016. For 2012 we screened all protocols manually for potential RCTs at each REC's archive, therefore the number of screened study protocols was higher than for 2016. For potential RCT protocols approved in 2016 we used the Business Administration System for Ethics Committees (BASEC) database to pre-screen study protocols for RCTs. Due to large numbers of RCT protocols in Zurich (KEK-ZH) and Basel (EKNZ), we took a random sample at each of these two RECs for feasibility reasons.

### 4.2 Characteristics of included Swiss RCT protocols

Overall, of the 400 included RCT protocols, 218 (54.5%) were non-industry-sponsored (i.e. investigator-initiated) and 182 (45.5%) were industry-sponsored RCT protocols (**Table 2**). The proportion of non-industry RCTs was 11% higher in 2016 compared to 2012 (89 [48.6%] for 2012 versus 129 [59.4%] for 2016). Overall, most RCTs were multicentre and parallel with individually randomized adults in the medical field with an average sample size of 200 participants (median total). Industry-sponsored protocols were, on average, larger, more frequently multicentre, placebo-controlled, testing a drug, and reporting to have had methodological support. Apart from the mentioned increase in the proportion of non-industry protocols, there were no substantial differences in RCT characteristics between 2012 and 2016.

Characteristics	2012			2016			Total
	Industry (N=94)	Non-industry (N=89)	Total (n=183)	Industry (N=88)	Non-industry (N=129)	Total (N=217)	N=400
<b>Planned target sample size, median (IQR)</b>	400 (184.5, 792.5)	136 (60, 382)	250 (91, 659)	403 (150, 652.5)	120 (60, 260)	180 (80, 490)	200 (80 - 550)
<b>Planned centers</b>							
Single centre	2 (2.1)	38 (42.7)	40 (21.9)	3 (3.5)	65 (50.4)	68 (31.3)	108 (27.0)
Multicentre	92 (97.9)	51 (57.3)	143 (78.1)	85 (96.5)	64 (49.6)	149 (68.7)	292 (73.0)
International	88 (95.7)	32 (62.7)	120 (83.9)	82 (96.5)	34 (53.1)	116 (77.85)	236 (80.8)
National	4 (4.3)	19 (37.3)	23 (16.1)	3 (3.4)	30 (46.9)	33 (22.15)	56 (19.2)
<b>Medical Field</b>							
Medical	81 (86.2)	62 (69.7)	143 (78.1)	76 (86.4)	92 (71.3)	168 (77.4)	311 (77.8)
Surgical	10 (10.6)	18 (20.2)	28 (15.3)	6 (6.8)	19 (14.7)	25 (11.5)	53 (13.2)
Paediatrics	2 (2.1)	7 (7.9)	9 (4.9)	5 (5.7)	6 (4.7)	11 (5.1)	20 (5.0)
Other	1 (1.1)	2 (2.2)	3 (1.6)	1 (1.1)	12 (9.3)	13 (6.0)	16 (4.0)
<b>Age Group</b>							
Adults	88 (93.6)	73 (83)	161 (88.5)	78 (88.6)	108 (83.7)	186 (85.7)	347 (87.0)
Only Elderly	3 (3.2)	5 (5.7)	8 (4.4)	3 (3.4)	7 (5.4)	10 (4.6)	18 (4.5)
Paediatrics	2 (2.1)	8 (9.1)	10 (5.5)	4 (4.5)	6 (4.7)	10 (4.6)	20 (5.0)
Adults and Elderly	1 (1.1)	0 (0)	1 (0.5)	1 (1.1)	4 (3.1)	5 (2.3)	6 (1.5)
Adults and children	0 (0)	2 (2.3)	2 (1.1)	2 (2.3)	4 (3.1)	6 (2.8)	8 (2.0)
<b>CTU or CRO support</b>	72 (76.6)	44 (49.4)	116 (63.4)	52 (59.1)	57 (44.2)	109 (50.2)	225 (56.3)
<b>Tertile of the year</b>							
1st tertile	34 (36.2)	36 (40.4)	70 (38.3)	29 (33.0)	47 (36.4)	76 (35)	146 (36.5)
2nd tertile	31 (33.0)	35 (39.3)	66 (36.1)	33 (37.5)	46 (35.7)	79 (36.4)	145 (36.3)
3rd tertile	29 (30.9)	18 (20.2)	47 (25.7)	26 (29.5)	36 (27.9)	62 (28.6)	109 (27.2)
<b>Placebo used</b>	47 (50.0)	24 (27.0)	71 (38.8)	52 (59.1)	35 (27.1)	87 (40.1)	158 (39.5)
<b>Unit of randomisation</b>							
Individuals	93 (98.9)	86 (96.7)	179 (97.8)	86 (97.8)	125 (96.9)	211 (97.3)	390 (97.5)
Clusters	0 (0.0)	1 (1.1)	1 (0.6)	1 (1.1)	3 (2.3)	4 (1.8)	5 (1.25)
Body parts	1 (1.1)	2 (2.2)	3 (1.6)	1 (1.1)	1 (0.8)	2 (0.9)	5 (1.25)
<b>Study design</b>							
Parallel	92 (97.8)	75 (84.3)	167 (91.3)	85 (96.6)	116 (89.9)	201 (92.6)	368 (92.00)
Crossover	1 (1.1)	9 (10.1)	10 (5.5)	1 (1.1)	10 (7.8)	11 (5.1)	21 (5.25)
Factorial	1 (1.1)	4 (4.5)	5 (2.7)	2 (2.3)	3 (2.3)	5 (2.3)	10 (2.50)
Other	0 (0.0)	1 (1.1)	1 (0.5)	0 (0.0)	0 (0.0)	0 (0)	1 (0.25)
<b>Study purpose</b>							
Superiority	75 (79.8)	69 (77.5)	144 (78.7)	73 (83.0)	102 (79.1)	175 (80.7)	319 (79.7)
Non-inferiority	16 (17.0)	13 (14.6)	29 (15.8)	13 (14.7)	22 (17.0)	35 (16.1)	64 (16.0)
Unclear	3 (3.2)	7 (7.9)	10 (5.5)	2 (2.3)	5 (3.9)	7 (3.2)	17 (4.3)
<b>Intervention</b>							
Drug	79 (84.04)	36 (40.45)	115 (62.84)	71 (80.68)	41 (31.78)	112 (51.61)	227 (56.75)
Other	15 (15.96)	53 (59.55)	68 (37.16)	17 (19.32)	88 (68.22)	105 (48.39)	173 (43.25)
<b>Research ethics committee</b>							
EC_BE	11 (11.7)	13 (14.6)	24 (13.1)	23 (26.1)	32 (24.8)	55 (25.3)	79 (19.75)
EC_EKNZ	30 (31.9)	23 (25.8)	53 (28.9)	17 (19.3)	24 (18.6)	41 (18.8)	94 (23.5)
EC_EKOS	6 (6.3)	9 (10.1)	15 (8.1)	8 (9.1)	6 (4.7)	14 (6.4)	29 (7.25)
EC_GE	2 (2.1)	2 (2.2)	4 (2.1)	6 (6.8)	20 (15.5)	26 (11.9)	30 (7.5)
EC_TI	3 (3.1)	4 (4.4)	7 (3.8)	3 (3.4)	5 (3.9)	8 (3.6)	15 (3.75)
EC_VD	11 (11.7)	18 (20.2)	29 (15.8)	9 (10.2)	17 (13.2)	26 (11.9)	55 (13.75)
EC_ZH	31 (32.9)	20 (22.4)	51 (27.8)	22 (25.0)	25 (19.4)	47 (21.6)	98 (24.5)

**Table 2:** Characteristics of all included 400 Swiss RCTs. Abbreviations: EC\_BE = Ethics Committee Bern, EC\_EKNZ = Ethics Committee Northwestern and Central Switzerland, EC\_EKOS = Ethics Committee Eastern Switzerland, EC\_GE = Ethics Committee Geneva, EC\_TI = Ethics Committee Ticino, EC\_VD = Ethics Committee Vaud, EC\_ZH = Ethics Committee Zurich.

### 4.3 Reporting of SPIRIT items in Swiss RCT protocols

#### 4.3.1 Overall description of results

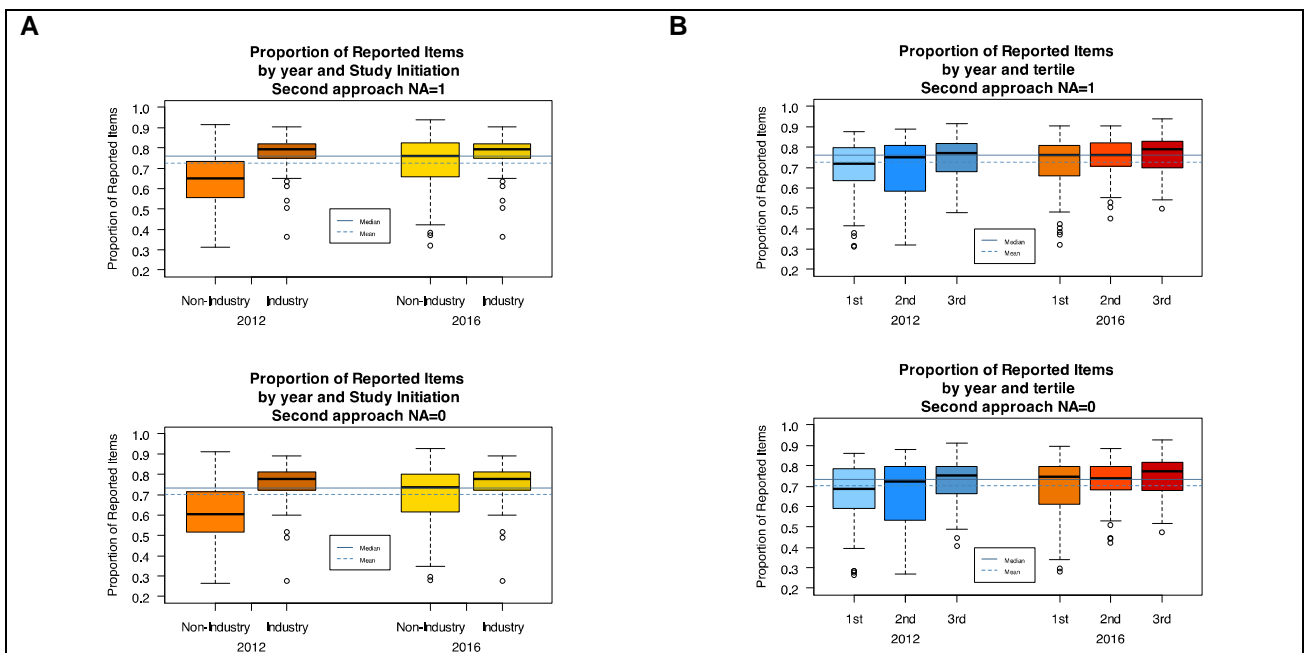
**Table 3** summarizes the adherence of reporting SPIRIT items per protocol based on our second approach (major item approach allowing for partial credit of individually met subitems or components of major SPIRIT items) with “not applicable” getting assigned a point.

Characteristic	2012						2016					
	Industry (N=94)		Non-industry (N=89)		Total 2012 (n=183)		Industry (N=88)		Non-industry (N=129)		Total 2016 (N=217)	
	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)
Frequency of items per protocol	26.13 (24.69-27.08)	25.54 (2.70)	21.42 (18.33-24.25)	20.99 (4.61)	24.46 (21.25-26.5)	23.33 (4.38)	25.42 (23.65-26.94)	24.87 (2.94)	25.17 (21.75-27.25)	24.22 (4.11)	25.17 (22.50-27.17)	24.48 (3.69)
Proportion of items per protocol	0.79 (0.75-0.82)	0.77 (0.082)	0.65 (0.56-0.74)	0.64 (0.14)	0.74 (0.64-0.80)	0.71 (0.13)	0.77 (0.72-0.82)	0.75 (0.09)	0.76 (0.66-0.83)	0.73 (0.13)	0.76 (0.69-0.82)	0.74 (0.11)

**Table 3:** Adherence to SPIRIT items in Swiss protocols stratified by year of approval and sponsorship.

Overall, the median proportion of reported items was similar in 2012 and 2016 (i.e. before and after introduction of the HRA): 74% in 2012 versus 76% in 2016. However, there was a moderate improvement, on average, of 11% in the proportion of met SPIRIT items for non-industry protocols (from a median of 65% in 2012 to a median of 76% in 2016 (**Figure 2A**)). Results based on the other two approaches including our sensitivity analyses were very similar and are presented in **AppendixTable 1**. It shows a range of median improvement in the proportion of met SPIRIT items between 8% and 13% depending on the applied approach.

Looking in more detail, how adherence to individual SPIRIT items or subitems changed between 2012 and 2016 stratified by sponsorship (**AppendixTable 2**), we found that the improvement in non-industry RCT protocols was due to an improvement in a large number of individual SPIRIT items and subitems. Overall, for 23 individual items or subitems the proportion of adherent protocols improved in non-industry RCTs by 10% or more. Among these 23 items and subitems were ones describing formal aspects (e.g. protocol version & date, or name & contact details of sponsor) as well as those describing methodological aspects (e.g. comparator choice explained, or allocation concealment mechanism). The largest improvements occurred with “description of process for making amendments” (SPIRIT item 25, +39%), “declaration of interests” (SPIRIT item 28, +32%), “name & contact details of sponsor” (SPIRIT item 5b, +21%), and “definition of analysis population” (SPIRIT item 20c, +20%). SPIRIT items with particularly low adherence in industry and non-industry sponsored protocols were “description of plans for granting access to full trial protocol” (SPIRIT item 31c, 4%) and “eligibility criteria for study centres” (SPIRIT item 10, 19%) in case of multicentre RCTs. We found no evidence for differences or trends in proportion of adherence to SPIRIT in terms of year tertiles (Jan-Apr; May-Aug; Sept-Dec) for 2012 and for 2016 (**Figure 2B**).



**Figure 2:** Proportion of reported SPIRIT items grouped by sponsorship and year. Figure 2A depicts the proportion of reported SPIRIT items grouped by tertiles of each year.



### 4.3.2 Results from multivariable regression analyses with Swiss RCT protocols

**Table 4** presents the results from our multivariable regression analyses using three different types of regression (linear, beta, and multilevel logistic). All different types of analyses including sensitivity analyses for our three different approaches yielded very similar results: Multicentre RCTs, RCTs with reported methodological support from CTUs or CROs, industry-sponsored RCTs, and RCTs approved in 2016 were independently and statistically significantly associated with better adherence to SPIRIT items. Whether the RCT protocol tested a drug or a non-drug intervention and the planned sample size were not consistently found to be independent predictors of adherence. The interaction term of year of approval and sponsorship turned out to be statistically significant as well as the likelihood ratio test providing evidence that the model including the interaction term describes the data significantly better than a model without the interaction term. This means that there is evidence for a differential improvement in the adherence to SPIRIT over time (2012 vs 2016) for industry-sponsored protocols and non-industry-sponsored protocols in the sense that there was an improvement for non-industry protocols but not for industry protocols (see Table 2). We did not find a statistically significant interaction for year of approval and CTU/CRO support, i.e. protocols with or without reported support from CTUs or CROs improved to a similar extent from 2012 to 2016.

Approach	Independent Variable	Linear Regression					Beta Regression					Log regression with Protocol as random effect				
		estimate	CI	p value	Chisq	p	estimate	CI	p value	Chisq	p	Odds Ratio	CI	p value	Chisq	p
First Approach NA 0	Sample size/1000	0.001	-0.007 – 0.008	0.845	-	-	1.004	0.974-1.035	0.809	-	-	1.004	0.973 – 1.035	0.821	-	-
	Multicenter	0.050	0.022 – 0.078	<.001	-	-	1.230	1.099-1.376	<.001	-	-	1.248	1.115 – 1.398	<.001	-	-
	CTU or CRO support	0.060	0.038 – 0.082	<.001	-	-	1.279	1.170-1.400	<.001	-	-	1.269	1.160 – 1.389	<.001	-	-
	Industry Initiation	0.045	0.020 – 0.071	<.001	-	-	1.192	1.075-1.321	0.001	-	-	1.204	1.086 – 1.336	<.001	-	-
	Year	0.053	0.032 – 0.074	<.001	-	-	1.245	1.144-1.355	<.001	-	-	1.225	1.126 – 1.334	<.001	-	-
	Drug intervention	0.033	0.008 – 0.057	0.009	-	-	1.147	1.040-1.265	0.006	-	-	1.170	1.060 – 1.291	0.002	-	-
	Interaction term	Initiation:Year	-0.095	-0.136 – -0.055	<.001	21.081	<.001	0.677	0.574-0.799	<.001	20.945	<.001	0.679	0.575 – 0.801	<.001	20.553
	CTU or CRO support:Year	-0.040	-0.082 – 0.002	0.062	3.555	0.059	0.840	0.708-0.996	0.044	4.038	0.045	0.844	0.711 – 1.001	0.052	3.780	0.052
First Approach NA 1	Sample size/1000	0.000	-0.009 – 0.008	0.924	-	-	1.000	0.965-1.036	0.998	-	-	0.998	0.962 – 1.036	0.926	-	-
	Multicenter	0.038	0.006 – 0.069	0.02	-	-	1.165	1.024-1.327	0.021	-	-	1.173	1.025 – 1.342	0.021	-	-
	CTU or CRO support	0.075	0.049 – 0.100	<.001	-	-	1.369	1.233-1.521	<.001	-	-	1.382	1.240 – 1.540	<.001	-	-
	Industry Initiation	0.047	0.018 – 0.076	0.002	-	-	1.208	1.071-1.363	0.002	-	-	1.223	1.080 – 1.386	0.001	-	-
	Year	0.061	0.037 – 0.084	<.001	-	-	1.293	1.172-1.427	<.001	-	-	1.299	1.174 – 1.439	<.001	-	-
	Drug intervention	0.026	-0.002 – 0.054	0.068	-	-	1.106	0.986-1.240	0.084	-	-	1.117	0.992 – 1.258	0.067	-	-
	Interaction term	Initiation:Year	-0.112	-0.159 – -0.066	<.001	22.616	<.001	0.628	0.518-0.761	<.001	21.804	<.001	0.620	0.508 – 0.756	<.001	21.843
	CTU or CRO support:Year	-0.027	-0.075 – 0.021	0.273	1.227	0.2681	0.908	0.745-1.107	0.338	0.918	0.338	0.906	0.738 – 1.112	0.346	0.887	0.346
Second Approach NA 0	Sample size/1000	0.001	-0.008 – 0.009	0.868	-	-	1.005	0.966-1.045	0.808	-	-	-	-	-	-	-
	Multicenter	0.040	0.009 – 0.072	0.012	-	-	1.174	1.022-1.348	0.023	-	-	-	-	-	-	-
	CTU or CRO support	0.071	0.046 – 0.096	<.001	-	-	1.393	1.244-1.559	<.001	-	-	-	-	-	-	-
	Industry Initiation	0.043	0.014 – 0.072	0.003	-	-	1.214	1.065-1.383	0.004	-	-	-	-	-	-	-
	Year	0.060	0.036 – 0.083	<.001	-	-	1.325	1.192-1.473	<.001	-	-	-	-	-	-	-
	Drug intervention	0.030	0.002 – 0.057	0.033	-	-	1.129	0.998-1.277	0.053	-	-	-	-	-	-	-
	Interaction term	Initiation:Year	-0.125	-0.170 – -0.080	<.001	29.003	<.001	0.569	0.463-0.699	<.001	27.592	<.001	-	-	-	-
	CTU or CRO support:Year	-0.043	-0.090 – 0.004	0.075	3.244	0.0717	0.861	0.696-1.063	0.164	1.929	0.165	-	-	-	-	-
Second Approach NA 1	Sample size/1000	0.000	-0.008 – 0.008	0.976	-	-	1.000	0.963-1.039	0.989	-	-	-	-	-	-	-
	Multicenter	0.032	0.002 – 0.061	0.034	-	-	1.135	0.991-1.299	0.067	-	-	-	-	-	-	-
	CTU or CRO support	0.067	0.044 – 0.091	<.001	-	-	1.386	1.242-1.548	<.001	-	-	-	-	-	-	-
	Industry Initiation	0.038	0.011 – 0.064	0.006	-	-	1.193	1.050-1.356	0.007	-	-	-	-	-	-	-
	Year	0.053	0.031 – 0.075	<.001	-	-	1.298	1.170-1.440	<.001	-	-	-	-	-	-	-
	Drug intervention	0.020	-0.006 – 0.045	0.128	-	-	1.078	0.956-1.217	0.220	-	-	-	-	-	-	-
	Interaction term	Initiation:Year	-0.114	-0.157 – -0.072	<.001	27.630	<.001	0.582	0.475-0.711	<.001	26.723	<.001	-	-	-	-
	CTU or CRO support:Year	-0.033	-0.078 – 0.011	0.142	2.2063	0.1374	0.897	0.729-1.103	0.302	1.065	0.302	-	-	-	-	-
Third Approach NA 0	Sample size/1000	0.003	-0.005 – 0.011	0.484	-	-	1.015	0.977-1.055	0.439	-	-	1.022	0.989 – 1.056	0.187	-	-
	Multicenter	0.045	0.015 – 0.075	0.004	-	-	1.206	1.056-1.378	0.006	-	-	1.305	1.159 – 1.470	<.001	-	-
	CTU or CRO support	0.067	0.042 – 0.091	<.001	-	-	1.356	1.217-1.51	<.001	-	-	1.289	1.172 – 1.418	<.001	-	-
	Industry Initiation	0.031	0.004 – 0.059	0.027	-	-	1.147	1.012-1.299	0.032	-	-	1.146	1.027 – 1.279	0.015	-	-
	Year	0.044	0.021 – 0.066	<.001	-	-	1.224	1.105-1.354	<.001	-	-	1.195	1.093 – 1.307	<.001	-	-
	Drug intervention	0.020	-0.006 – 0.047	0.136	-	-	1.074	0.954-1.209	0.236	-	-	1.108	0.998 – 1.230	0.054	-	-
	Interaction term	Initiation:Year	-0.112	-0.156 – -0.068	<.001	24.988	<.001	0.603	0.495-0.735	<.001	24.250	<.001	0.657	0.552 – 0.782	<.001	21.762
	CTU or CRO support:Year	-0.026	-0.072 – 0.020	0.265	1.2684	0.260	0.921	0.751-1.129	0.428	0.628	0.428	0.882	0.737 – 1.057	0.174	1.840	0.175
Third Approach NA 1	Sample size/1000	0.002	-0.005 – 0.009	0.614	-	-	1.010	0.975-1.047	0.582	-	-	1.009	0.974 – 1.046	0.613	-	-
	Multicenter	0.028	0.001 – 0.055	0.039	-	-	1.127	0.995-1.277	0.059	-	-	1.145	1.004 – 1.306	0.044	-	-
	CTU or CRO support	0.060	0.038 – 0.081	<.001	-	-	1.336	1.207-1.478	<.001	-	-	1.356	1.220 – 1.508	<.001	-	-
	Industry Initiation	0.028	0.003 – 0.052	0.027	-	-	1.138	1.012-1.279	0.031	-	-	1.143	1.012 – 1.291	0.032	-	-
	Year	0.037	0.017 – 0.057	<.001	-	-	1.200	1.092-1.32	<.001	-	-	1.202	1.089 – 1.328	<.001	-	-
	Drug intervention	0.015	-0.009 – 0.038	0.214	-	-	1.049	0.939-1.172	0.393	-	-	1.067	0.950 – 1.199	0.27	-	-
	Interaction term	Initiation:Year	-0.100	-0.138 – -0.061	<.001	25.197	<.001	0.619	0.514-0.744	<.001	25.007	<.001	0.605	0.499 – 0.734	<.001	25.33
	CTU or CRO support:Year	-0.017	-0.057 – 0.024	0.421	0.660	0.4165	0.958	0.792-1.159	0.660	0.194	0.660	0.951	0.779 – 1.162	0.625	0.24	0.626

**Table 4:** Results from all regression analyses (linear regression, beta regression, and multilevel logistic regression) on 400 Swiss RCT protocols considering the three different approaches of counting the reported SPIRIT items and the two ways of considering items allowing for a “not applicable” option. Abbreviations: NA, not applicable; CTU, clinical trial unit; CRO, clinical research organisation.

#### 4.4 Registration of Swiss RCT protocols

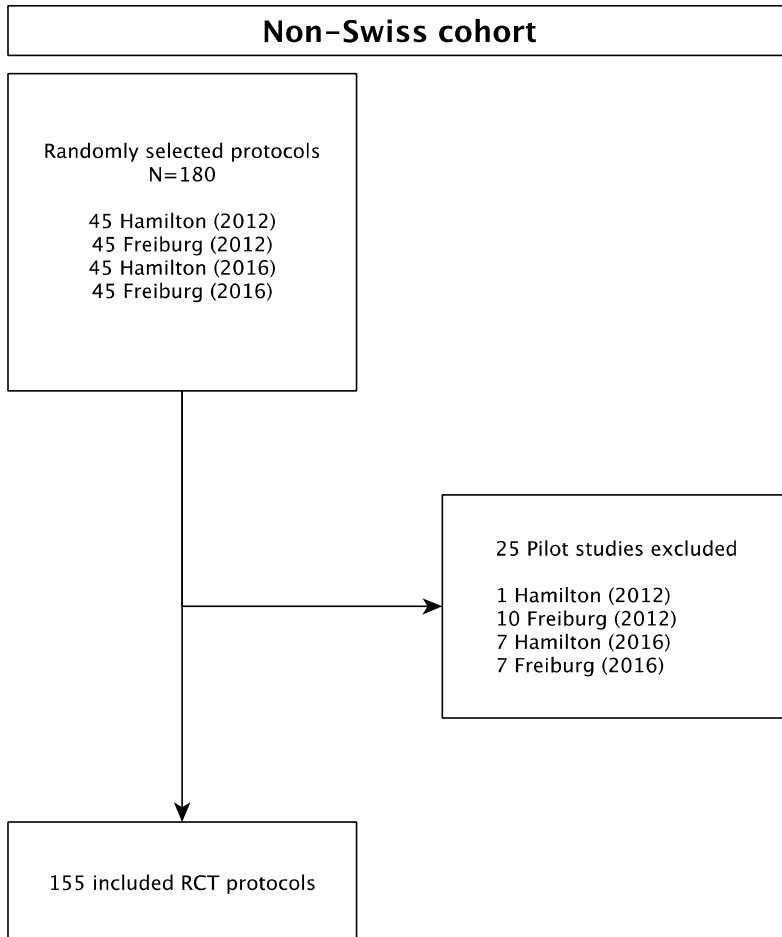
We found that industry-sponsored protocols were more frequently registered and more frequently prospectively registered than non-industry-sponsored protocols (e.g. for 2012: proportion of prospectively registered protocols of 89%, 95% CI, 82%-94% for industry-sponsored protocols vs 69%, 95% CI, 58%-77% for non-industry-sponsored protocols; **Table 5**). There was no evidence that the proportion of registered protocols or the proportion of prospectively registered protocols changed for industry-sponsored or non-industry-sponsored Swiss protocols from 2012 to 2016.

Characteristic	2012			2016			Total N=400
	Industry (N=94)	Non-industry (N=89)	Total 2012 (N=183)	Industry (N=88)	Non-industry (N=129)	Total 2016 (N=217)	
Registered in an International Registry	90 (95.74 (89.56- 98.33))	74 (83.15 (74.04- 89.51))	164 (89.62 (84.35- 93.25))	81 (92.05 (84.48- 96.09))	103 (79.85 (72.11- 85.86))	184 (84.79 (79.41- 88.96))	348 (87 (83.35- 89.95))
Not registered in an International Registry	4 (4.26 (1.67- 10.44))	15 (16.85 (10.49- 25.96))	19 (10.38 (6.75- 15.65))	7 (7.96 (3.91- 15.52))	26 (20.16 (14.14- 27.89))	33 (15.21 (11.04- 20.59))	52 (13 (10.05- 16.65))
Prospectively registered	84 (89.36 (81.52- 94.12))	61 (68.54 (58.30- 77.25))	145 (79.24 (72.79- 84.48))	77 (87.50 (78.99- 92.88))	92 (71.32 (62.99- 78.42))	169 (77.88 (71.90- 82.89))	314 (78.50 (74.21- 82.25))
Not prospectively registered	10 (10.64 (5.88- 18.49))	28 (31.46 (22.75- 41.71))	38 (20.77 (15.52- 27.21))	11 (12.50 (7.13- 21.01))	37 (28.68 (21.58- 37.01))	48 (22.12 (17.11- 28.10))	86 (21.50 (17.76- 25.79))

**Table 5:** Registration of Swiss RCTs in trial registries. Numbers are frequencies (proportion (95% confidence interval)).

#### 4.5 Differences between Swiss and non-Swiss protocols

Overall, we included a random sample of 155 eligible non-Swiss RCT protocols that were approved in 2012 or 2016 by RECs in Freiburg (Germany) or Hamilton (Canada) in our analyses (**Figure 3**). The characteristics of non-Swiss protocols are presented in **Table 6**. In general, the characteristics of the non-Swiss RCTs were similar to the Swiss RCTs with mostly multicentre and parallel RCTs with individually randomized adults in the medical field with a somewhat larger average sample size of 282 participants (median total; Swiss RCT protocols had a median of 200). Like with protocols approved in Switzerland industry-sponsored protocols approved outside of Switzerland were, on average, larger, more frequently multicentre, placebo-controlled, testing a drug, and reporting to have had methodological support than non-industry-sponsored protocols.



**Figure 3:** Flow diagram of non-Swiss protocol selection for the years 2012 and 2016

Characteristic	2012			2016			Total N=155
	Industry (N=46)	Non-industry (N=33)	Total 2012 (N=79)	Industry (N=41)	Non-industry (N=35)	Total 2016 (N=76)	
Planned target sample size , median (IQR)	485 (158.5- 1109)	198 (100-600)	400 (128.5-775)	270 (150- 444)	196 (107- 402)	205 (129.5- 417)	282 (128.5- 600)
Planned centers							
Single centre	0 (0.0)	7 (21.2)	7 (8.9)	1 ( 2.4)	14 (40.0)	15 (19.7)	22 (14.2)
Multicentre	46 (100.0)	26 (78.8)	72 (90.1)	40 ( 97.6)	21 (60.0)	61 (80.3)	133 (85.8)
International	40 (87.0)	14 (53.8)	54 (75.0)	35 (87.5)	9 (42.9)	44 (72.1)	98 (73.7)
National	6 (13.0)	12 (46.2)	18 (25.0)	5 (12.5)	12 (57.1)	17 (27.9)	35 (26.3)
Medical Field							
Medical	44 (95.7)	18 (54.5)	62 (78.5)	38 (92.7)	26 (74.3)	64 (84.2)	126 (81.3)
Surgical	1 (2.2)	9 (27.3)	10 (12.6)	0 (0.0)	6 (17.1)	6 (7.9)	16 (10.3)
Paediatrics	1 (2.2)	5 (15.2)	6 (7.6)	2 (4.9)	3 (8.6)	5 (6.6)	11 (7.1)
Other	0 (0.0)	1 (3.0)	1 (1.3)	1 (2.4) (NR)	0 (0.0)	1 (1.3)	2 (1.3)
Age Group							
Adults	42 (91.3)	23 (69.7)	65 (82.3)	34 (82.9)	25 (71.4)	59 (77.6)	124 (80.0)
Only Elderly	1 (2.2)	2 (6.1)	3 (3.8)	0 (0.0)	3 (8.6)	3 (3.9)	6 (3.9)
Paediatrics	0 (0.0)	4 (12.0)	4 (5.1)	3 (7.3)	7 (20.0)	10 (13.2)	14 (9.0)
Adults and Elderly	0 (0.0)	2 (6.1)	2 (2.5)	1 (2.4)	0 (0.0)	1 (1.4)	3 (1.9)
Adults and children	3 (6.5)	2 (6.1)	5 (6.3)	3 (7.3)	0 (0.0)	3 (3.9)	8 (5.2)
CTU or CRO support	23 (50)	12 (36)	35 (43.2)	22 (54)	11 (31)	33 (43.4)	68 (43.9)
Tertile	1 missing						
1st tertile	17 (37.0)	13 (39.4)	30 (38.5)	11 (26.8)	13 (37.1)	24 (31.6)	54 (35.1)
2nd tertile	15 (32.6)	11 (33.3)	26 (33.3)	13 (31.7)	12 (34.3)	25 (32.9)	51 (33.1)
3rd tertile	13 (28.3)	9 (27.3)	22 (28.2)	17 (41.5)	10 (28.6)	27 (35.5)	49 (31.8)
Placebo used	31 (67)	9 (27)	40 (50.6)	24 (59)	9 (26)	33 (43.4)	73 (47.1)
Unit of randomisation							
Individuals	46 (100.0)	30 (90.9)	76 (96.2)	40 (97.6)	35 (100.0)	75 (98.7)	151 (97.4)
Clusters	0 (0.0)	3 (9.1)	3 (3.8)	0 (0.0)	0 (0.0)	0 (0.0)	3 (1.9)
Body parts	0 (0.0)	0 (0.0)	0 (0.0)	1 (2.4)	0 (0.0)	1 (1.3)	1 (0.7)
Study design							
Parallel	45 (97.8)	30 (90.9)	75 (95.0)	40 (97.6)	32 (91.4)	72 (94.7)	147 (94.8)
Crossover	1 (2.2)	1 (3.0)	2 (2.5)	1 (2.4)	2 (5.7)	3 (4.0)	5 (3.2)
Factorial	0 (0.0)	2 (6.1)	2 (2.5)	0 (0.0)	1 (2.9)	1 (1.3)	3 (2.0)
Study purpose							
Superiority	38 (82.6)	25 (75.8)	63 (79.7)	33 (80.5)	31 (88.6)	64 (84.2)	127 (81.9)
Non-inferiority	7 (15.2)	6 (18.2)	13 (16.5)	7 (17.1)	1 (2.9)	8 (10.5)	21 (13.6)
Unclear	1 (2.2)	2 (6.1)	3 (3.8)	1 (2.4)	3 (8.6)	4 (5.3)	7 (4.5)
Intervention							
Drug	38 (82.61)	13 (39.39)	51 (64.56)	36 (87.80)	16 (45.71)	52 (68.42)	103 (66.45)
Other	8 (17.39)	20 (60.61)	28 (35.44)	5 (12.20)	19 (54.29)	24 (31.58)	52 (33.55)
Research ethics committee							
Canada (Hamilton)	22 (47.8)	22 (66.7)	44 (55.7)	14 (34.2)	24 (68.6)	38 (50.0)	82 (52.9)
Germany (Freiburg)	24 (52.2)	11 (33.3)	35 (44.3)	27 (65.8)	11 (31.4)	38 (50.0)	73 (47.1)

**Table 6:** Characteristics of all included 155 non-Swiss RCTs.

Overall, the median proportion of reported items in non-Swiss protocol was practically the same in 2012 and 2016: 69% in 2012 versus 69% in 2016 (**Table 7**). However, there was a modest improvement, on average, of 3% in the proportion of met SPIRIT items for non-industry protocols (from a median of 59% in 2012 to a median of 62% in 2016). Adherence to SPIRIT was, on average, lower in non-Swiss protocols when compared to Swiss protocols (compare **Table 3** with **Table 7**, and **first part Table 8**). In the subgroup of non-Swiss protocols, we found evidence for a differential improvement in adherence to SPIRIT in non-industry-sponsored protocols from 2012 to 2016, too, while the adherence did not change in the subgroup of industry-sponsored protocols (**Table 7, middle part Table 8**).

Characteristic	2012						2016					
	Industry (N=46)		Non-industry (N=33)		Total 2012 (N=79)		Industry (N=41)		Non-industry (N=35)		Total 2016 (N=76)	
	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)
Frequency of items per protocol	23.92 (22.33-25.02)	23.41 (2.31)	19.33 (17.42-22.92)	19.27 (4.246)	22.83 (19.42-24.42)	21.66 (3.82)	23.92 (22.50-25.25)	23.24 (3.76)	20.50 (17.71-23.00)	20.52 (4.54)	22.67 (19.73-24.94)	21.98 (4.33)
Proportion of items per protocol	0.73 (0.68-0.76)	0.71 (0.07)	0.59 (0.53-0.69 )	0.58 (0.13)	0.69 (0.59-0.74)	0.66 (0.13)	0.73 (0.68-0.77)	0.70 (0.11)	0.62 (0.54-0.70)	0.62 (0.14)	0.69 (0.60-0.76)	0.67 (0.13)

**Table 7:** Adherence to SPIRIT items in non-Swiss protocols stratified by year of approval and sponsorship.

Approach	Independent Variable	Linear Regression			Likelihood ratio		Beta Regression			Likelihood ratio	
		Estimate	CI	p value	Chisq	p	Odds Ratio	CI	p value	Chisq	p
Protocols from Switzerland, Germany (Freiburg) and Canada (Hamilton) (n=555)											
Second Approach NA 0	Sample size/1000	0.004	0.000–0.009	0.048	-	-	1.02	1.000-1.041	0.048	-	-
	Multicenter	0.060	0.033–0.087	<.001	-	-	1.296	1.152-1.458	<.001	-	-
	CTU or CRO support	0.064	0.043–0.084	<.001	-	-	1.345	1.227-1.475	<.001	-	-
	Industry Initiation	0.039	0.015–0.062	0.001	-	-	1.163	1.046-1.292	0.005	-	-
	Year	0.052	0.033–0.071	<.001	-	-	1.282	1.175-1.399	<.001	-	-
	Swiss Cohort	0.074	0.053–0.096	<.001	-	-	1.406	1.276-1.549	<.001	-	-
	Drug Intervention	0.044	0.021–0.067	<.001	-	-	1.214	1.096-1.344	<.001	-	-
Interaction term	Initiation:Year	-0.113	-0.151–0.076	<.001	34.846	<.001	0.596	0.503-0.706	<.001	34.604	<.001
	CTU or CRO support:Year	-0.022	-0.060–0.017	0.264	1.269	0.26	0.944	0.794-1.122	0.515	0.423	0.516
Second Approach NA 1	Sample size/1000	0.004	-0.001–0.008	0.088	-	-	1.012	0.990-1.034	0.308	-	-
	Multicenter	0.049	0.024–0.074	<.001	-	-	1.314	1.151-1.498	<.001	-	-
	CTU or CRO support	0.060	0.041–0.080	<.001	-	-	1.434	1.295-1.589	<.001	-	-
	Industry Initiation	0.034	0.012–0.056	0.002	-	-	0.990	0.880-1.115	0.871	-	-
	Year	0.046	0.028–0.064	<.001	-	-	1.343	1.218-1.480	<.001	-	-
	Swiss Cohort	0.070	0.050–0.091	<.001	-	-	1.219	1.094-1.359	<.001	-	-
	Drug Intervention	0.034	0.012–0.055	0.002	-	-	1.251	1.116-1.403	<.001	-	-
Interaction term	Initiation:Year	-0.101	-0.136–0.066	<.001	31.642	<.001	0.548	0.454-0.662	<.001	37.700	<.001
	CTU or CRO support:Year	-0.013	-0.049–0.023	0.479	0.509	0.476	1.110	0.915-1.346	0.290	1.109	0.292
Only protocols from Germany (Freiburg) and Canada (Hamilton) (n=155)											
Second Approach NA 0	Sample size/1000	0.004	-0.000–0.009	0.058	-	-	1.020	0.999-1.042	0.067	-	-
	Multicenter	0.145	0.094–0.196	<.001	-	-	1.885	1.512-2.351	<.001	-	-
	CTU or CRO support	0.054	0.021–0.086	0.001	-	-	1.297	1.121-1.501	<.001	-	-
	Industry Initiation	0.030	-0.007–0.066	0.107	-	-	1.072	0.911-1.262	0.402	-	-
	Year	0.032	-0.000–0.064	0.053	-	-	1.189	1.029-1.373	0.019	-	-
	Drug Intervention	0.082	0.044–0.121	<.001	-	-	1.441	1.216-1.707	<.001	-	-
Interaction term	Initiation:Year	-0.078	-0.140–0.016	0.014	6.345	0.012	0.673	0.510-0.888	0.005	7.655	0.006
	CTU or CRO support:Year	-0.001	-0.063–0.062	0.980	<.001	0.979	1.068	0.805-1.416	0.649	0.207	0.649
Second Approach NA 1	Sample size/1000	0.004	-0.001–0.008	0.100	-	-	1.016	0.987-1.046	0.281	-	-
	Multicenter	0.123	0.074–0.171	<.001	-	-	2.137	1.556-2.935	<.001	-	-
	CTU or CRO support	0.051	0.020–0.082	0.001	-	-	1.569	1.270-1.938	<.001	-	-
	Industry Initiation	0.026	-0.008–0.061	0.134	-	-	0.713	0.561-0.906	0.006	-	-
	Year	0.027	-0.004–0.057	0.084	-	-	1.404	1.141-1.728	0.001	-	-
	Drug Intervention	0.070	0.033–0.107	<.001	-	-	1.643	1.286-2.098	<.001	-	-
Interaction term	Initiation:Year	-0.063	-0.122–0.003	0.039	4.491	0.034	0.458	0.308-0.681	<.001	14.370	<.001
	CTU or CRO support:Year	0.008	-0.051–0.068	0.779	0.083	0.773	1.767	1.181-2.642	0.006	7.404	0.007
Only Non- Industry Initiated Protocols from Switzerland, Germany (Freiburg) and Canada (Hamilton) (n=286)											
Second Approach NA 0	Sample size/1000	0.001	-0.013–0.015	0.902	-	-	1.000	0.941-1.062	0.987	-	-
	Multicenter	0.051	0.018–0.084	0.003	-	-	1.254	1.087-1.446	0.002	-	-
	CTU or CRO support	0.083	0.050–0.116	<.001	-	-	1.454	1.259-1.679	<.001	-	-
	Year	0.105	0.074–0.135	<.001	-	-	1.594	1.397-1.818	<.001	-	-
	Swiss Cohort	0.088	0.051–0.124	<.001	-	-	1.434	1.229-1.675	<.001	-	-
	Drug Intervention	0.042	0.009–0.075	0.014	-	-	1.199	1.037-1.385	0.014	-	-
Interaction term	CTU or CRO support:Year	-0.016	-0.077–0.046	0.621	0.251	0.616	0.991	0.757-1.297	0.947	0.005	0.947
	Swiss Cohort:Year	0.062	-0.008–0.133	0.084	3.087	0.079	1.268	0.939-1.710	0.121	2.407	0.121
Second Approach NA 1	Sample size/1000	0.001	-0.013–0.014	0.921	-	-	0.959	0.895-1.028	0.233	-	-
	Multicenter	0.041	0.010–0.072	0.01	-	-	1.262	1.070-1.487	0.006	-	-
	CTU or CRO support	0.078	0.047–0.109	<.001	-	-	1.613	1.365-1.906	<.001	-	-
	Year	0.093	0.064–0.122	<.001	-	-	1.716	1.474-2.999	<.001	-	-
	Swiss Cohort	0.083	0.049–0.117	<.001	-	-	1.108	0.925-1.326	0.266	-	-
	Drug Intervention	0.030	-0.001–0.062	0.056	-	-	1.260	1.065-1.490	0.007	-	-
Interaction term	CTU or CRO support:Year	-0.007	-0.065–0.051	0.818	0.054	0.816	1.300	0.951-1.770	0.101	2.616	0.106
	Swiss Cohort:Year	0.061	-0.006–0.127	0.073	3.322	0.068	0.844	0.594-1.199	0.343	0.910	0.34

**Table 8:** Results from regression analyses (linear and beta regression) considering the second approach of counting the reported SPIRIT items and the two ways of considering items allowing for a “not applicable” option based on all 555 Swiss and non-Swiss protocols (upper part), based on 155 non-Swiss protocols only (middle part), and based on all 286 non-industry-sponsored Swiss and non-Swiss protocols (bottom part). Abbreviations: NA, not applicable; CTU, clinical trial unit; CRO, clinical research organisation.

Since the improvement in adherence to SPIRIT items appeared more pronounced with non-industry-sponsored RCT protocols approved in Switzerland (improvement of about 11%) than the improvement in non-industry-sponsored protocols approved outside of Switzerland (improvement of about 4%), we tested for an interaction between year of approval and geographic area of approval (in Switzerland vs outside of Switzerland) (**bottom part of Table 8**). There was a trend towards a more pronounced improvement in Swiss non-industry-sponsored protocols, which was, however, not statistically significant.

Like with Swiss RCT protocols the proportion of registered and the proportion of prospectively registered industry-sponsored protocols were larger than the respective proportions in non-industry-sponsored protocols approved outside of Switzerland (e.g. for 2012: proportion of prospectively registered protocols of 91%, 95% CI, 80%-97% for industry-sponsored protocols vs 73%, 95% CI, 56%-85% for non-industry-sponsored protocols; **Table 9**). There was some evidence that the proportion of prospectively registered protocols increased in the subgroup of non-industry-sponsored non-Swiss protocols from 73%, 95% CI, 56%-85% in 2012 to 83%, 95% CI, 67%-92% in 2016. Such a change was not seen for non-industry-sponsored Swiss protocols nor for industry-sponsored protocols (see Tables 5 and 9).

Characteristic	2012			2016			Total N=155
	Industry (N=46)	Non-industry (N=33)	Total 2012 (N=79)	Industry (N=41)	Non-industry (N=35)	Total 2016 (N=76)	
Registered in an International Registry	46 (100 (92.29- 100))	28 (84.85 (69.08-93.35))	74 (93.67 (86.03-97.27))	40 (97.56 (87.41-99.88))	29 (82.86 (67.32-91.90))	69 (90.79 (82.19-95.47))	143 (92.26 (86.96-95.52))
Not registered in an International Registry	0 (0 (0.00- 7.71))	5 (15.15 (6.65- 30.92))	5 (6.33 (2.73- 13.98))	1 (2.44 (0.13- 12.60))	6 (17.14 (8.10- 32.68))	7 (9.21 (4.53- 17.81))	12 (7.74 (4.48- 13.04))
Prospectively registered	42 (91.30 (79.68-96.57))	24 (72.73 (55.78-84.93))	66 (83.54 (73.85-90.12))	38 (92.68 (80.57-97.48))	29 (82.86 (67.32-91.90))	67 (88.16 (79.00-93.64))	133 (85.81 (79.45-90.44))
Not prospectively registered	4 (8.70 (3.43- 20.33))	9 (27.27 (15.07-44.22))	13 (16.46 (9.88- 26.15))	3 (7.32 (2.52- 19.43))	6 (17.14 (8.10- 32.68))	9 (11.84 (6.36- 21.00))	22 (14.19 (9.56- 20.56))

**Table 9:** Registration of non-Swiss RCTs in trial registries. Numbers are frequencies (proportion (95% confidence interval))

## 5 DISCUSSION

### 5.1 Summary of findings

In this before-after study we investigated the completeness of 183 RCT protocols that were approved by Swiss RECs in 2012 (two years before the introduction of the HRA) and 217 RCT protocols that were approved by Swiss RECs in 2016 (two years after the introduction of the HRA) using the SPIRIT-checklist. Overall, we did not find a relevant difference in the proportion of reported items according to SPIRIT between RCT protocols from 2012 (median 74%) and 2016 (median 76%), however, we found a significant improvement for the subgroup of non-industry-sponsored protocols (i.e. investigator-initiated RCTs) from a median proportion of 65% in 2012 to a median proportion of reported SPIRIT items of 76% in 2016. This improvement in non-industry RCT protocols was due to an improvement in the reporting of a large number of individual SPIRIT items and subitems with 23 individual items improving by 10% or more in the proportion of adherent protocols. The proportion of reported SPIRIT items in industry-sponsored protocols remained stable over time (median of 79% in 2012 vs 77% in 2016). This subgroup effect was confirmed to be independent of the planned sample size of RCTs, reported support from a CTU or CRO, drug vs non-drug intervention, and centre status (single- vs multicentre RCTs). We found that the following RCT characteristics were significantly and independently associated with lower adherence to SPIRIT: single centre RCT, no reported support from CTU or CRO, non-industry-sponsored (i.e. investigator-initiated), and approved in 2012. When we additionally investigated 155 RCT protocols that were approved by RECs outside of Switzerland (Freiburg, Germany, and Hamilton, Canada) in 2012 and 2016, we found a subgroup effect with an improvement of non-industry-sponsored protocols from 2012 to 2016, too, while the proportion of reported SPIRIT items remained about the same for industry-sponsored protocols. We did not find a subgroup effect for protocols with and without methodological support from CTUs or CROs. Although not being statistically significant, there was a trend for a larger improvement of Swiss compared to non-Swiss RCT protocols in terms of reported SPIRIT items. With respect to the proportion of registered and, in particular, prospectively registered protocols we found that industry-sponsored protocols were more frequently registered and more frequently prospectively registered than non-industry-sponsored protocols approved in Switzerland or outside of Switzerland in 2012 and 2016. There was no evidence for an increase in the proportion of registered or prospectively registered protocols approved in Switzerland.

### 5.2 Strengths and limitations

Strengths of our study include full access to the protocols and associated documents of all included RCTs approved by Swiss RECs or RECs in Freiburg and Hamilton in 2012 and 2016. We used standardized methods and procedures for data extraction and protocol assessment at all sites (i.e. RECs) and involved only trained methodologists in this process. To minimize chance associations, we considered only a limited number of variables in the statistical models. Our results proved robust in sensitivity analyses applying alternative assumptions and statistical approaches. The fact that practically all Swiss RECs participated in this study strengthens the representativeness of our data for Switzerland and the additional consideration of German and Canadian RCT protocols allowed for an international perspective to better differentiate between developments exclusively happening in Switzerland and broader/international trends.

Our study has several limitations. First, we used single data extraction and assessment for 70% of protocols approved by Swiss RECs and the REC in Hamilton in 2016 for feasibility reasons, thereby potentially increasing extraction errors. However, we used pre-piloted extraction forms with detailed written instructions and conducted calibration exercises with all data extractors. More than 95% of included protocols approved in 2012 and protocols approved in Freiburg in 2016 were extracted and assessed in duplicate. Second, we used a convenience sample of two RECs outside of Switzerland (Freiburg in Germany, Hamilton in Canada). We cannot say whether they are representative for other RECs in these or other countries; to our knowledge they are not in any way particular. The originally planned inclusion of additional RCT protocols approved by RECs in the UK was not possible, because data extraction and protocol assessment could not be completed in time due to a lack of methodologically trained data extractors. Third, since we used RCT protocols for this study that had already been approved by RECs, SPIRIT items such as “research ethics approval” and “consent forms provided” were always fulfilled and did, therefore, not contribute to discriminate more accurate/complete protocols from less accurate/complete protocols. Fourth, although we had adequate statistical power to detect even modest differences in the adherence to SPIRIT for non-industry sponsored protocols approved outside of Switzerland, we might have lacked statistical power in the interaction analysis of improvement in adherence for

non-industry-sponsored RCT protocols approved in Switzerland and outside of Switzerland. Finally, our assumption that the adherence to SPIRIT as a measure for the accuracy/completeness of reporting of RCT protocols indeed reflects the “quality of RCTs” or the “quality of clinical research in general” is based on scientific reasoning and common sense rather than empirical evidence. Further research is necessary to examine the association of accuracy/completeness of RCT protocols with risks for premature discontinuation or non-publication of RCTs (see 5.4. below).

### **5.3 Comparison of the results with similar studies**

There are only few studies in the literature that have used (19) or planned to use (20,21) the SPIRIT checklist as a tool to assess the completeness of trial protocols. The completed study by Kyte et al. investigated 75 RCT protocols from the UK National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme on the reporting of patient-reported outcomes and the association with general protocol completeness according to the SPIRIT checklist (19). They found that these 75 non-industry-sponsored UK RCT protocols from 2012 and 2013 included a mean of 63% of SPIRIT checklist items, which is very similar to our findings for non-industry-sponsored RCT protocols from 2012 in Switzerland with a mean of 64% (Table 3). Apart from the present study, we are not aware of any evaluation studies on the completeness of RCT protocols including industry-sponsored protocols or examining protocol completeness over time.

Various studies examined trial registration and, in particular, prospective trial registration based on published RCTs (22–26). They found prospective or “adequate” registration rates in the ballpark of 30% to 45% for RCTs published between 2008 and 2014 with varying definitions of “adequate registration” (e.g. “registered before the end of the trial” (22) or “before enrolment of first patient” (23)). We are aware of only a single study that determined a trial registration rate based on RCT protocols as we did (18). Chan et al. examined 53 industry-sponsored and 60 non-industry sponsored RCT protocols (total of 113 protocols) approved at two Finnish RECs in 2007. Using the same definition of “prospectively registered” as we did (i.e. within one month after the trial start or enrolment of the first patient) they found 61% of RCTs to be prospectively registered. The article did not report separate prospective registration rates for industry-sponsored and non-industry sponsored protocols. Our findings of 79% of prospectively registered Swiss protocols and 83% of prospectively registered non-Swiss protocols in 2012 suggests a somewhat continuing international improvement in prospective registration in the decade following the statement of the International Committee of Medical Journal Editors (ICMJE) in 2004 (27).

### **5.4 Conclusions and future directions**

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

We interpret our findings as evidence for an international trend of a moderate improvement in the completeness of non-industry-sponsored (i.e. investigator-initiated) RCT protocols 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, and minimizing reporting biases in the scientific community; and efforts to teach RCT methodology to clinician scientists in under- and postgraduate courses. 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 from swissethics for study protocols 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.

Comprehensive and well-reported study protocols are essential to ensure the safety and well-being of study participants, data validity, successful study conduct, and credibility of results, particularly in the case of RCTs. Incomplete protocols jeopardize the clinical research process at all stages with potentially harmful consequences for patients, decision-makers in health care, the scientific community, and society as a whole. Further empirical research is necessary to investigate whether improvements in protocol adherence to SPIRIT recommendations reduce, for instance, the proportion of prematurely discontinued RCTs or the proportion of RCTs that were not published in peer-reviewed journals or without published results in trial registries.



## 6 ACKNOWLEDGMENT

We would like to thank all the presidents and staff of the Swiss RECs, the REC in Freiburg (Germany), and the REC in Hamilton (Canada) as well as Swissethics for their support and collaboration in this study. We sincerely acknowledge the contributions from Belinda von Niederhäusern, Benjamin Speich, Elena Ojeda Ruiz, Stefan Schandelmaier, Lars G. Hemkens, Kimberly McCord, Alain Nordmann, Alain Amstutz, Christiane Pauli-Magnus, Yuki Tomonaga, Matthias Schwenkglens, Anette Blümle, Karin Bischoff, Katharina Kunzweiler, Laura Rehner, Jason W. Busse, Dominik Mertz, Jacqueline Wong, Ngai Chow, and Patrick Ji Ho Hong in extracting data and assessing included RCT protocols. We are very grateful to Ayodele Odutayo, Sally Hopewell, An-Wen Chan, and Doug Altman for their advice and input to the study design, and their help with developing the data extraction forms. We are also grateful to Ramon Saccilotto for adapting the data extraction tool “squakeo.org” to our needs and for all his trouble shooting whenever needed. Finally, we thank the Swiss Federal Office of Public Health for supporting and funding this project.

## 7 REFERENCES

1. WMA DECLARATION OF HELSINKI – ETHICAL PRINCIPLES FOR MEDICAL RESEARCH INVOLVING HUMAN SUBJECTS [Internet]. Available from: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>
2. Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database Syst Rev*. 2011 Jan 19;(1):MR000031.
3. Tetzlaff JM, Chan A-W, Kitchen J, Sampson M, Tricco AC, Moher D. Guidelines for randomized clinical trial protocol content: a systematic review. *Syst Rev*. 2012 Jan;1(1):43.
4. Scharf O, Colevas AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. *J Clin Oncol Off J Am Soc Clin Oncol*. 2006 Aug 20;24(24):3933–8.
5. Pildal J, Chan A-W, Hróbjartsson A, Forfang E, Altman DG, Gøtzsche PC. Comparison of descriptions of allocation concealment in trial protocols and the published reports: cohort study. *BMJ*. 2005 May;330(7499):1049.
6. Hróbjartsson A, Pildal J, Chan A-W, Haahr MT, Altman DG, Gøtzsche PC. Reporting on blinding in trial protocols and corresponding publications was often inadequate but rarely contradictory. *J Clin Epidemiol*. 2009 Sep;62(9):967–73.
7. Chan A-W, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA J Am Med Assoc*. 2004 May;291(20):2457–65.
8. Al-Marzouki S, Roberts I, Evans S, Marshall T. Selective reporting in clinical trials: analysis of trial protocols accepted by *The Lancet*. *Lancet*. 2008 Jul;372(9634):201.
9. Smyth RMD, Kirkham JJ, Jacoby A, Altman DG, Gamble C, Williamson PR. Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. *Br Med J*. 2011 Jan;342:c7153.
10. Chan A-W, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, et al. Increasing value and reducing waste: addressing inaccessible research. *The Lancet*. 2014 Jan;6736(13):1–10.
11. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. *Ann Intern Med*. 2013 Feb;158(3):200–7.
12. Decullier E. Fate of biomedical research protocols and publication bias in France: retrospective cohort study. *BMJ*. 2005 Jul;331(7507):19.
13. Getz KA, Zuckerman R, Cropp AB, Hindle AL, Krauss R, Kaitin KI. Measuring the Incidence, Causes, and Repercussions of Protocol Amendments. *Drug Inf J*. 2011 May;45(3):265–75.
14. Kasenda B, Schandelmaier S, Sun X, von Elm E, You J, Blümle A, et al. Subgroup analyses in randomised controlled trials: cohort study on trial protocols and journal publications. *BMJ*. 2014 Jul;349(jul16 1):g4539.
15. Chan A-W, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ*. 2013 Jan;346(jan08 15):e7586–e7586.
16. Kasenda B, von Elm E, You J, Blümle A, Tomonaga Y, Saccilotto R, et al. Prevalence, characteristics, and publication of discontinued randomized trials. *JAMA J Am Med Assoc*. 2014 März;311(10):1045–51.
17. Ferrari S, Cribari-Neto F. Beta Regression for Modelling Rates and Proportions. *J Appl Stat*. 2004 Aug;31(7):799–815.
18. Chan A-W, Pello A, Kitchen J, Axentiev A, Virtanen JI, Liu A, et al. Association of Trial Registration With Reporting of Primary Outcomes in Protocols and Publications. *JAMA*. 2017 07;318(17):1709–11.
19. Kyte D, Duffy H, Fletcher B, Gheorghe A, Mercieca-Bebber R, King M, et al. Systematic evaluation of the patient-reported outcome (PRO) content of clinical trial protocols. *PloS One*. 2014;9(10):e110229.
20. Madden K, Arseneau E, Evaniew N, Smith CS, Thabane L. Reporting of planned statistical methods in published surgical randomised trial protocols: a protocol for a methodological systematic review. *BMJ Open*. 2016 03;6(6):e011188.
21. Thabane A, Dennis BB, Gajic-Veljanoski O, Paul J, Thabane L. Reporting quality of stepped wedge design randomized trials: a systematic review protocol. *Clin Epidemiol*. 2016;8:261–6.
22. Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA J Am Med Assoc*. 2009 Sep;302(9):977–84.
23. Scott A, Rucklidge JJ, Mulder RT. Is Mandatory Prospective Trial Registration Working to Prevent Publication of Unregistered Trials and Selective Outcome Reporting? An Observational Study of Five Psychiatry Journals That Mandate Prospective Clinical Trial Registration. *PloS One*. 2015;10(8):e0133718.
24. Harriman SL, Patel J. When are clinical trials registered? An analysis of prospective versus retrospective registration. *Trials*. 2016 Apr 15;17:187.
25. Farquhar CM, Showell MG, Showell EAE, Beetham P, Baak N, Mourad S, et al. Clinical trial registration was not an indicator for low risk of bias. *J Clin Epidemiol*. 2017 Apr;84:47–53.
26. Zarin DA, Tse T, Williams RJ, Rajakannan T. Update on Trial Registration 11 Years after the ICMJE Policy Was Established. *N Engl J Med*. 2017 26;376(4):383–91.
27. DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA*. 2004 Sep 15;292(11):1363–4.

## 8 APPENDIX

**Appendix Table 1:** Adherence to the SPIRIT items per protocol stratified by year of approval and sponsorship.

Characteristic	2012						2016					
	Industry (n=94 (51.4))		Non-industry (n=89 (48.6))		Total 2012 (n=183)		Industry (n=88 (40.6))		Non-industry (n=129 (59.4))		Total 2016 (n=217)	
	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)
<b>First Approach Na=0</b>												
Frequency of items reported	19 (17-20.75)	18.7 (2.87)	13 (11-17)	13.76 (4.36)	17 (13-20)	16.3 (4.42)	18.5 (17-20)	18.27 (3.29)	17 (14-19)	16.45 (3.88)	18 (15-20)	17.19 (3.75)
Proportion of items reported	0.5758 (0.5313-0.6335)	0.5698 (0.0858)	0.4194 (0.3333-0.5313)	0.4231 (0.1321)	0.5313 (0.4063-0.6061)	0.4985 (0.1327)	0.5691 (0.5152-0.6108)	0.5593 (0.0978)	0.5313 (0.4242-0.5938)	0.5097 (0.1180)	0.5455 (0.4545-0.6061)	0.5298 (0.1127)
<b>First Approach Na=1</b>												
Frequency of items reported	22 (20.25-24)	21.85 (3.17)	17 (15-20)	16.79 (4.86)	20 (16-23)	19.4 (4.79)	22 (20.25-24)	21.85 (3.17)	17 (15-20)	16.79 (4.86)	21 (18-24)	20.69 (4.11)
Proportion of items reported	0.6667 (0.6136-0.7273)	0.6622 (0.0961)	0.5152 (0.4545-0.6061)	0.509 (0.1474)	0.6061 (0.4848-0.6969)	0.5877 (0.1453)	0.6667 (0.6061-0.6970)	0.6488 (0.1009)	0.6364 (0.5152-0.7273)	0.6122 (0.1367)	0.6364 (0.5455-0.7273)	0.6270 (0.1245)
<b>Second Approach Na=0</b>												
Frequency of items reported	25.38 (23.66-26.56)	24.83 (2.96)	19.75 (16.63-22.92)	19.47 (5.00)	23.58 (19.40-25.94)	22.23 (4.88)	24.54 (22.88-25.89)	24.04 (3.27)	23.92 (19.88-25.75)	22.79 (4.34)	24.25 (21.58-25.75)	23.29 (3.98)
Proportion of items reported	0.7778 (0.7246-0.8106)	0.7569 (0.0885)	0.6042 (0.5172-0.7135)	0.5988 (0.1515)	0.7222 (0.5992-0.7917)	0.6799 (0.1462)	0.7507 (0.7031-0.7958)	0.7360 (0.0968)	0.7374 (0.6151-0.8000)	0.7065 (0.1334)	0.7475 (0.6692-0.7995)	0.7185 (0.1205)
<b>Second Approach Na=1</b>												
Frequency of items reported	26.13 (24.69-27.08)	25.54 (2.70)	21.42 (18.33-24.25)	20.99 (4.61)	24.46 (21.25-26.5)	23.33 (4.38)	25.42 (23.65-26.94)	24.87 (2.94)	25.17 (21.75-27.25)	24.22 (4.11)	25.17 (22.50-27.17)	24.48 (3.69)
Proportion of items reported	0.7917 (0.7481-0.8207)	0.7738 (0.0818)	0.6490 (0.5556-0.7348)	0.6361 (0.1398)	0.7412 (0.6439-0.8030)	0.7069 (0.1328)	0.7702 (0.7165-0.8163)	0.7535 (0.0891)	0.7626 (0.6591-0.8258)	0.7340 (0.1246)	0.7626 (0.6818-0.8232)	0.7419 (0.1118)
<b>Third Approach Na=0</b>												
Frequency of items reported	44.00 (41.00-47.00)	43.33 (5.18)	35.00 (30.00-39.00)	34.54 (8.41)	41 (34.50-45.00)	39.06 (8.21)	42.50 (39.75-46.00)	41.82 (6.16)	40.00 (35.00-45.00)	39.40 (7.80)	42 (37.00-45.00)	40.38 (7.26)
Proportion of items reported	0.7586 (0.7143-0.8025)	0.7457 (0.0846)	0.6182 (0.5283-0.7000)	0.6086 (0.1408)	0.7119 (0.6087-0.7787)	0.6790 (0.1340)	0.7333 (0.6820-0.7782)	0.7160 (0.0978)	0.7170 (0.6140-0.7895)	0.6948 (0.1317)	0.7241 (0.6491-0.7797)	0.7034 (0.1193)
<b>Third Approach Na=1</b>												
Frequency of items reported	30.00 (27.00-33.00)	29.26 (4.76)	23.00 (20.00-26.00)	22.89 (7.90)	27.00 (24.00-30.00)	25.67 (7.44)	28.50 (25.00-32.00)	27.50 (5.43)	28.00 (24.00-32.00)	26.77 (7.34)	28.00 (24.00-32.00)	27.07 (6.63)
Proportion of items reported	0.7813 (0.7344-0.8125)	0.7696 (0.0744)	0.6719 (0.5781-0.7344)	0.6545 (0.1235)	0.7344 (0.6563-0.7969)	0.7136 (0.1163)	0.7578 (0.7031-0.7969)	0.7422 (0.0849)	0.7500 (0.6563-0.8125)	0.7307 (0.1147)	0.7500 (0.6875-0.7969)	0.7354 (0.1036)

**Appendix Table 2: Reporting of each SPIRIT item stratified by year and sponsorship.**

Variable	2012										2016										
	Spirit Item Number	Industry (n=94 (51.4))			Non-Industry (n=89 (48.6))			Total 2012 (n=183)				Industry (n=88 (40.6))			Non-Industry (n=129 (59.4))				Total 2016 (n=217)		
		Yes	Not applicable	No	Yes	Not applicable	No	Yes	Not applicable	No	Yes	Not applicable	No	Yes	Not applicable	No	Yes	Not applicable	No		
Basic study design in Title	1	71 (75.53)	NA	23 (24.47)	26 (29.21)	NA	63 (70.79)	97 (53.01)	NA	86 (46.99)	70 (79.55)	NA	18 (20.45)	46 (35.66)	NA	83 (64.34)	116 (53.46)	NA	101 (46.54)		
Trial registration	2	78 (82.98)	NA	16 (17.02)	56 (62.92)	NA	33 (37.08)	111 (60.66)	NA	72 (39.34)	75 (85.23)	NA	13 (14.77)	99 (76.74)	NA	30 (23.26)	174 (80.18)	NA	43 (19.82)		
Protocol version, number and date	3	91 (96.81)	NA	3 (3.19)	77 (86.52)	NA	12 (13.48)	168 (91.8)	NA	15 (8.2)	88 (100)	NA	0 (0)	128 (99.22)	NA	1 (0.78)	216 (99.54)	NA	1 (0.46)		
Funding sources	4	90 (95.74)	NA	4 (4.26)	59 (66.29)	NA	30 (33.71)	149 (81.42)	NA	34 (18.58)	82 (93.18)	NA	6 (6.82)	105 (81.4)	NA	24 (18.6)	187 (86.18)	NA	30 (13.82)		
Names of protocol contributors/authors	5a	22 (23.40)	NA	72 (76.60)	29 (32.58)	NA	60 (67.42)	51 (27.87)	NA	132 (72.13)	16 (18.18)	NA	72 (81.82)	19 (14.73)	NA	110 (85.27)	35 (16.13)	NA	182 (83.87)		
Name and contact details of sponsor	5b	82 (87.23)	NA	12 (12.77)	63 (70.79)	NA	26 (29.21)	145 (79.23)	NA	38 (20.77)	52 (59.09)	NA	36 (40.91)	118 (91.47)	NA	11 (8.53)	170 (78.34)	NA	47 (21.66)		
Role of sponsor and funder in trial	5c	84 (89.36)	NA	10 (10.64)	32 (35.96)	NA	57 (64.04)	116 (63.39)	NA	67 (36.61)	34 (38.64)	NA	54 (61.36)	36 (27.91)	NA	93 (72.09)	70 (32.26)	NA	147 (67.74)		
Steering Committee General Membership and Role	5d	23 (24.47)	62 (65.96)	9 (9.57)	24 (26.97)	59 (66.29)	6 (6.74)	47 (25.68)	121 (66.12)	15 (8.2)	15 (17.05)	58 (65.91)	15 (17.05)	33 (25.58)	94 (72.87)	2 (1.55)	48 (22.12)	152 (70.05)	17 (7.83)		
Research question described and justified	6a	17 (18.09)	NA	77 (81.91)	22 (24.72)	NA	67 (75.28)	39 (21.31)	NA	144 (78.69)	17 (19.32)	NA	71 (80.68)	41 (31.78)	NA	88 (68.22)	58 (26.73)	NA	159 (73.27)		
Comparator choice explained	6b	78 (82.98)	NA	16 (17.02)	65 (73.03)	NA	24 (26.97)	143 (78.14)	NA	40 (21.86)	74 (84.09)	NA	14 (15.91)	111 (86.05)	NA	18 (13.95)	185 (85.25)	NA	32 (14.75)		
Specific objectives described	7	91 (96.81)	NA	3 (3.19)	79 (88.76)	NA	10 (11.24)	170 (92.9)	NA	13 (7.1)	84 (95.45)	NA	4 (4.55)	120 (93.02)	NA	9 (6.98)	204 (94.01)	NA	13 (5.99)		
Trial design described	8	86 (91.49)	NA	8 (8.51)	58 (65.17)	NA	31 (34.83)	144 (78.69)	NA	39 (21.31)	83 (94.32)	NA	5 (5.68)	101 (78.29)	NA	28 (21.71)	184 (84.79)	NA	33 (15.21)		
Countries where data will be collected listed	9	63 (67.02)	NA	31 (32.98)	75 (84.27)	NA	14 (15.73)	138 (75.41)	NA	45 (24.59)	16 (18.18)	NA	72 (81.82)	112 (86.82)	NA	17 (13.18)	228 (58.99)	NA	89 (41.01)		
Eligibility criteria for trial participants	10	94 (100)	NA	0 (0)	88 (98.88)	NA	1 (1.12)	182 (99.45)	NA	1 (0.55)	88 (100)	NA	0 (0)	129 (100)	NA	0 (0)	217 (100)	NA	0 (0)		
Eligibility criteria for study centres and who will perform the intervention	10	13 (13.83)	1 (1.06)	80 (85.11)	13 (14.61)	36 (40.45)	40 (44.94)	26 (14.21)	37 (20.22)	120 (65.77)	7 (7.95)	2 (2.27)	79 (89.77)	24 (18.6)	60 (46.51)	45 (34.88)	31 (14.29)	62 (28.57)	124 (57.14)		
Individuals administering interventions (non-drug)	10	10 (10.64)	78 (82.98)	6 (6.38)	31 (34.83)	37 (41.57)	21 (23.6)	41 (22.4)	115 (62.84)	27 (14.75)	11 (12.5)	67 (76.14)	10 (11.36)	56 (43.41)	47 (36.43)	26 (20.16)	67 (30.88)	114 (52.53)	36 (16.59)		
Generic Name Dose and Schedule of intervention	11a	80 (85.10)	12 (12.77)	2 (2.13)	41 (46.07)	47 (52.81)	1 (1.12)	121 (66.12)	59 (32.24)	3 (1.64)	74 (84.09)	14 (15.91)	0 (0)	46 (35.66)	82 (63.57)	1 (0.78)	120 (55.3)	96 (44.24)	1 (0.46)		
Setting of intervention administration	11a	7 (7.45)	78 (82.98)	9 (9.57)	38 (42.7)	37 (41.57)	14 (15.73)	45 (24.59)	115 (62.84)	23 (12.57)	12 (13.64)	67 (76.14)	9 (10.23)	70 (54.26)	45 (34.88)	14 (10.85)	82 (37.79)	112 (51.61)	23 (10.6)		
Criteria for modifications of interventions	11b	70 (74.47)	10 (10.64)	14 (14.89)	41 (46.07)	23 (25.84)	25 (28.09)	111 (60.66)	33 (18.03)	39 (21.31)	75 (85.23)	4 (4.55)	9 (10.23)	73 (56.59)	30 (23.26)	26 (20.16)	148 (68.2)	34 (15.67)	35 (16.13)		
Strategies to improve or monitor adherence	11c	51 (54.26)	35 (37.23)	8 (8.51)	21 (23.6)	52 (58.43)	16 (17.98)	72 (39.34)	87 (47.54)	24 (13.11)	54 (61.36)	25 (28.41)	9 (10.23)	49 (37.98)	61 (47.29)	19 (14.73)	103 (47.27)	86 (39.63)	28 (12.9)		
Permitted concomitant care	11d	89 (94.68)	NA	5 (5.32)	43 (48.31)	NA	46 (51.69)	132 (72.13)	NA	51 (27.87)	83 (94.32)	NA	5 (5.68)	97 (75.91)	NA	32 (24.81)	180 (82.95)	37 (17.05)	187 (85.9)		
Primary Outcome: Specific measurement variable	12	94 (100)	NA	0 (0)	85 (95.51)	NA	4 (4.49)	179 (97.81)	NA	4 (2.19)	87 (98.86)	NA	1 (1.14)	119 (92.25)	NA	10 (7.75)	206 (94.93)	NA	11 (5.07)		
Primary Outcome: Analysis metric	12	88 (93.62)	1 (1.06)	5 (5.32)	72 (80.9)	0 (0)	17 (19.1)	160 (87.43)	1 (0.55)	22 (12.02)	82 (93.18)	1 (1.14)	5 (5.68)	102 (79.07)	0 (0)	27 (20.93)	184 (84.79)	1 (0.46)	32 (14.75)		
Primary Outcomes: time point of measurement	12	60 (63.83)	29 (30.85)	5 (5.32)	63 (70.79)	16 (17.98)	10 (11.24)	123 (67.21)	45 (24.59)	15 (8.2)	67 (76.14)	17 (19.32)	4 (4.55)	102 (79.07)	12 (9.3)	15 (11.63)	169 (77.88)	29 (13.36)	19 (8.76)		
Participant timeline	13	93 (98.94)	NA	1 (1.06)	85 (95.51)	NA	4 (4.49)	178 (97.27)	NA	5 (2.73)	87 (98.86)	NA	1 (1.14)	126 (97.67)	NA	3 (2.33)	213 (98.16)	NA	4 (1.84)		
Sample size: Estimated number	14	94 (100)	NA	0 (0)	88 (98.88)	NA	1 (1.12)	182 (99.45)	NA	1 (0.55)	87 (98.86)	NA	1 (1.14)	128 (99.22)	NA	1 (0.78)	215 (99.08)	NA	2 (0.92)		
Sample size: Outcome used for sample size calculation	14	87 (92.55)	4 (4.26)	3 (3.19)	79 (88.76)	2 (2.25)	8 (8.99)	166 (90.71)	6 (3.28)	11 (6.01)	82 (93.18)	3 (3.41)	3 (3.41)	108 (83.72)	7 (5.43)	14 (10.85)	190 (87.56)	10 (4.61)	17 (7.83)		
Sample size: Assumed values for outcome	14	84 (89.36)	3 (3.19)	7 (7.45)	62 (69.66)	4 (4.49)	23 (25.84)	146 (79.78)	7 (3.83)	30 (16.39)	66 (75.00)	3 (3.41)	19 (21.59)	84 (65.12)	7 (5.43)	38 (29.46)	150 (69.12)	10 (4.61)	57 (26.27)		
Sample size: Alpha value	14	86 (91.49)	4 (4.26)	4 (4.26)	77 (86.52)	2 (2.25)	10 (11.24)	163 (89.07)	6 (3.28)	14 (7.65)	83 (94.32)	3 (3.41)	2 (2.27)	108 (83.72)	7 (5.43)	14 (10.85)	191 (88.02)	10 (4.61)	16 (7.37)		
Sample size: Statistical Power	14	89 (94.68)	4 (4.26)	1 (1.06)	81 (91.01)	2 (2.25)	6 (6.74)	170 (92.9)	6 (3.28)	7 (3.83)	83 (94.32)	3 (3.41)	2 (2.27)	113 (87.6)	7 (5.43)	9 (6.98)	196 (90.32)	10 (4.61)	11 (5.07)		
Sample size: Rationale sample size if not derived statistically	14	4 (4.26)	89 (94.68)	1 (1.06)	0 (0)	81 (91.01)	8 (8.99)	4 (2.19)	170 (92.9)	9 (4.92)	1 (1.14)	85 (96.59)	2 (2.27)	3 (2.33)	121 (93.8)	5 (3.88)	4 (1.84)	206 (94.93)	7 (3.23)		
Location of participant recruitment	15	15 (15.96)	NA	79 (84.04)	60 (67.42)	NA	29 (32.58)	75 (40.98)	NA	108 (59.02)	12 (13.64)	NA	76 (86.36)	97 (75.19)	NA	32 (24.81)	109 (50.23)	NA	108 (49.77)		
Person(s) who will recruit participants	15	28 (29.79)	NA	66 (70.21)	38 (42.7)	NA	51 (57.3)	66 (36.07)	NA	117 (63.93)	25 (28.41)	NA	63 (71.59)	77 (59.69)	NA	52 (40.31)	102 (47)	NA	115 (53)		
Expected recruitment rate	15	31 (32.98)	NA	63 (67.02)	36 (40.45)	NA	53 (59.55)	67 (36.61)	NA	116 (63.39)	13 (14.77)	NA	75 (85.23)	24 (18.6)	NA	105 (81.4)	37 (17.05)	NA	180 (82.95)		
Method for generation of random sequence	16a	58 (61.70)	NA	36 (38.30)	46 (51.69)	NA	43 (48.31)	104 (56.83)	NA	79 (43.17)	49 (55.68)	NA	39 (44.32)	85 (65.89)	NA	44 (34.11)	134 (61.75)	NA	83 (38.25)		
Allocation concealment mechanism	16b	80 (85.11)	7 (7.45)	7 (7.45)	58 (65.17)	3 (3.37)	28 (31.46)	138 (75.41)	10 (5.46)	35 (19.13)	76 (86.36)	0 (0)	12 (13.64)	102 (79.07)	3 (2.33)	24 (18.6)	178 (82.03)	3 (1.38)	36 (16.59)		
Person who will enroll/assign participants	16c	42 (44.68)	NA	52 (55.32)	30 (33.71)	NA	59 (66.29)	72 (39.34)	NA	111 (60.66)	41 (46.59)	NA	47 (53.41)	61 (47.29)	NA	68 (52.71)	102 (47)	NA	115 (53)		
Blinding status of participants	17a	91 (96.81)	NA	3 (3.19)	76 (85.39)	NA	13 (14.61)	167 (91.26)	NA	16 (8.74)	85 (96.59)	NA	3 (3.41)	115 (89.15)	NA	14 (10.85)	200 (92.17)	NA	17 (7.83)		
Blinding status of care providers	17a	92 (97.87)	NA	2 (2.13)	77 (86.52)	NA	12 (13.48)	169 (92.35)	NA	14 (7.65)	85 (96.59)	NA	3 (3.41)	116 (89.92)	NA	13 (10.08)	201 (92.63)	NA	16 (7.37)		
Blinding status of outcome assessors	17a	74 (78.72)	NA	20 (21.28)	50 (56.18)	NA	39 (43.82)	124 (67.76)	NA	59 (32.24)	56 (63.64)	NA	32 (36.36)	77 (59.69)	NA	52 (40.31)	133 (61.29)	NA	84 (38.71)		
Conditions when unblinding is permissible	17b	89 (62.77)	30 (31.91)	5 (5.32)	19 (21.35)	50 (56.18)	20 (22.47)	78 (42.62)	80 (43.72)	25 (13.66)	54 (61.36)	24 (27.27)	10 (11.36)	44 (34.11)	67 (51.94)	18 (13.95)	98 (45.16)	91 (41.94)	28 (12.9)		
Personnel who will collect data	18a	40 (42.55)	NA	54 (57.45)	39 (43.82)	NA	50 (56.18)	79 (43.17)	NA	104 (56.83)	43 (48.86)	NA	45 (51.14)	68 (52.71)	NA	61 (47.29)	111 (51.15)	NA	106 (48.85)		
Strategies to promote participant retention and complete follow-up	18b	61 (64.89)	NA	33 (35.11)	27 (30.34)	NA	62 (69.66)	88 (48.09)	NA	95 (51.91)	59 (67.05)	NA	29 (32.95)	48 (37.21)	NA	81 (62.79)	107 (49.31)	NA	110 (50.69)		
Data entry and coding	19	81 (86.17)	NA	13 (13.83)	54 (60.67)	NA	35 (39.31)	135 (73.77)	NA	48 (26.23)	69 (78.41)	NA	19 (21.59)	95 (73.64)	NA	34 (26.36)	164 (75.58)	NA	53 (24.42)		
Main analysis for primary outcome	20a	89 (94.68)	NA	5 (5.32)	68 (76.4)	NA	21 (23.6)	157 (85.79)	NA	26 (14.21)	84 (95.45)	NA	4 (4.55)	102 (79.07)	NA	27 (20.93)	186 (85.71)	NA	31 (14.29)		
Definition of subgroup categories	20b	40 (42.55)	43 (45.74)	11 (11.7)	12 (13.48)	61 (68.54)	16 (17.98)	52 (28.42)	104 (56.83)	27 (14.75)	24 (27.27)	45 (51.14)	19 (21.59)	15 (11.63)	101 (78.29)	13 (10.08)	39 (17.97)	146 (67.28)	32 (14.75)		
Definition of analysis population	20c	84 (89.36)	NA	10 (10.64)	36 (40.45)	NA	53 (59.55)	120 (65.57)	NA	63 (34.43)	79 (89.77)	NA	9 (10.23)	78 (60.47)	NA	51 (39.53)	157 (72.35)	NA	60 (27.65)		
DMC is planned or why it is not planned	21a	70 (74.47)	NA	24 (25.53)	33 (37.08)	NA	56 (62.92)	103 (56.28)	NA	80 (43.72)	67 (76.14)	NA	21 (23.86)	45 (34.88)	NA	84 (65.12)	112 (51.61)	NA	105 (48.39)		
Who has authority to stop the trial	21b	76 (80.85)	NA	18 (19.15)	52 (58.43)	NA	37 (41.57)</														

**Appendix Table 3:** Adherence to SPIRIT items in Swiss protocols stratified by year of approval and methodological support (yes vs no)

Characteristic	2012						2016					
	CTU or CRO support (N=116)		No CTU or CRO support (N=67)		Total 2012 (N=183)		CTU or CRO support (N=109)		No CTU or CRO support (N=108)		Total 2016 (N=217)	
	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)
<b>Second Approach 1*</b>												
Frequency of items per protocol	24.58 (21.92-26.33)	23.82 (3.524)	19.75 (16.07-24.75)	19.47 (5.638)	23.58 (19.40-25.94)	22.23 (4.88)	25.25 (23.33-26.67)	24.55 (3.334)	22.97 (19.24-25.17)	22.02 (4.185)	24.25 (21.58-25.75)	23.29 (3.98)
Proportion of items per protocol	0.7551 (0.6679-0.7986)	0.73 (0.1054)	0.6010 (0.4945-0.7500)	0.60 (0.1695)	0.72 (0.5992-0.7917)	0.6799 (0.1462)	0.7740 (0.7151-0.8182)	0.76 (0.1005)	0.71 (0.5996-0.7788)	0.68 (0.1273)	0.7475 (0.6692-0.7995)	0.7185 (0.1205)
<b>Second Approach 2*</b>												
Frequency of items per protocol	25.33 (22.81-27.08)	24.69 (3.249)	21.75 (17.58-25.58)	20.97 (5.066)	24.46 (21.25-26.5)	23.33 (4.38)	26.33 (24.33-27.67)	25.64 (3.049)	24.12 (21.23-26.27)	23.31 (3.915)	25.17 (22.50-27.17)	24.48 (3.69)
Proportion of items per protocol	0.77 (0.6913-0.8207)	0.75 (0.09846)	0.6591 (0.5328-0.7753)	0.64 (0.15351)	0.74 (0.6439-0.8030)	0.71 (0.1328)	0.80 (0.7374-0.8384)	0.78 (0.09240)	0.73 (0.6433-0.7961)	0.7064 (0.11863)	0.7626 (0.6818-0.8232)	0.74 (0.1118)

**Appendix Table 4:** Adherence to SPIRIT items in Swiss protocols stratified by year of approval and centre status (single vs multi-centre)

Characteristic	2012						2016					
	Multicenter (N=143)		Singlecenter (N=40)		Total 2012 (N=183)		Multicenter (N=149)		Singlecenter (N=68)		Total 2016 (N=217)	
	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)	median (IQR)	mean (SD)
<b>Second Approach 1*</b>												
Frequency of items per protocol	24.58 (21.40-26.25)	23.42 (4.157)	18.84 (14.14-21.03)	17.96 (4.925)	23.58 (19.40-25.94)	22.23 (4.88)	24.58 (22.42-26.25)	24.02 (3.247)	23.25 (18.27-25.43)	21.71(4.904)	24.25 (21.58-25.75)	23.29 (3.98)
Proportion of items per protocol	0.7551 (0.6526-0.7980)	0.7144 (0.12449)	0.5931 (0.4419-0.6497)	0.5568 (0.15305)	0.7222 (0.5992-0.7917)	0.6799 (0.1462)	0.7500 (0.6869-0.8030)	0.7375 (0.09736)	0.7266 (0.5708-0.7889)	0.6766 (0.15263)	0.7475 (0.6692-0.7995)	0.7185 (0.1205)
<b>Second Approach 2*</b>												
Frequency of items per protocol	25.33 (22.75-26.79)	24.29 (3.766)	20.62 (17.11-22.93)	19.90 (4.746)	24.46 (21.25-26.5)	23.33 (4.38)	25.50 (23.33-27.17)	24.98 (3.020)	25.08 (20.08-26.88)	23.38 (4.684)	25.17 (22.50-27.17)	24.48 (3.69)
Proportion of items per protocol	0.77 (0.6894-0.8119)	0.7359 (0.11413)	0.6250 (0.5186-0.6948)	0.6029 (0.14381)	0.7412 (0.6439-0.8030)	0.7069 (0.1328)	0.7727 (0.7071-0.8232)	0.7571 (0.09151)	0.7601 (0.6086-0.8144)	0.7086 (0.14193)	0.7626 (0.6818-0.8232)	0.7419 (0.1118)

**Appendix Table 5:** Adherence to SPIRIT items in non-industry sponsored protocols stratified by year of approval and geographic area (Switzerland vs other)

Characteristic	2012									2016								
	Swiss (N=89)			Non-Swiss (N=33)			Total 2012 (N=122)			Swiss (N=129)			Non-Swiss (N=35)			Total 2016 (N=164)		
	median (IQR)	mean	SD	median (IQR)	mean	SD	median (IQR)	mean	SD	median (IQR)	mean	SD	median (IQR)	mean	SD	median (IQR)	mean	SD
<b>Second Approach NA = 0</b>																		
Frequency of items reported per protocol	19.75 (16.63-22.92)	19.47	5.000	17.67 (15.75-21.08)	17.58	4.787	19.21 (16.25-22.18)	18.96	4.995	23.92 (19.88-25.75)	22.79	4.336	18.92 (15.55-21.54)	18.96	4.584	22.62 (18.67-25.50)	21.97	4.650
Proportion of items reported per protocol	0.6042 (0.5172-0.7135)	0.5987	0.1515	0.5469 (0.4773-0.6616)	0.5439	0.1468	0.5974 (0.5045-0.6850)	0.5839	0.1516	0.7374 (0.6151-0.8000)	0.7065	0.1334	0.5807 (0.4859-0.6629)	0.5906	0.1437	0.7090 (0.5806-0.7923)	0.6817	0.1433
<b>Second Approach NA = 1</b>																		
Frequency of items reported per protocol	21.42 (18.33-24.25)	20.99	4.613	19.33 (17.42-22.92)	19.27	4.246	21.12 (18.06-23.67)	20.53	4.565	25.17 (21.75-27.25)	24.22	4.113	20.50 (17.71-23.00)	20.52	4.541	24.21 (20.35-26.94)	23.43	4.461
Proportion of items reported per protocol	0.6490 (0.5556-0.7348)	0.6361	0.1398	0.5859 (0.5278-0.6944)	0.5839	0.1287	0.6402 (0.5473-0.7172)	0.6220	0.1383	0.7626 (0.6591-0.8258)	0.7340	0.1246	0.6212 (0.5366-0.6970)	0.6218	0.1376	0.7336 (0.6168-0.8163)	0.7101	0.1352