Data Monitoring Committees and clinical trials: from scientific justification to organisation

Clara Locher, Silvy Laporte, Peggy Derambure, Olivier Chassany, Cécile Girault, Alix Avakiantz, Claire Bahans, Dominique Deplanque, Pierre Fustier, Anne Françoise Germe, et al.

THERAPIES

Giens Workshops 2023/Organisational et regulatory aspects

Data Monitoring Committees and clinical trials: from scientific justification to organisation*

Clara Locher[a,*], Silvy Laporte[b], Peggy Derambure[c], Olivier Chassany[d], Cécile Girault[e], Alix Avakiantz[f,1], Claire Bahans[g,1], Dominique Deplanque[h,1], Pierre Fustier[i,1], Anne-Françoise Germe[j,1], Behrouz Kassaï[k,1], Louis Lacoste[l,1], Nadine Petitpain[m,1], Matthieu Roustit[n,1], Tabassome Simon[o,1], Cécile Train[p,1], Michel Cucherat[q,1]

[a] *CHU Rennes, Inserm, Centre d'investigation clinique de Rennes (CIC1414), service de pharmacologie clinique, Institut de recherche en santé, environnement et travail (Irset), UMR S 1085, 35000 Rennes, France*

[b] *Univ. Jean Monnet Saint-Etienne, UMR 1059 Inserm, URC/Pharmacologie Clinique, CHU Saint-Etienne, 42055 Saint-Etienne, France*

[c] *AbbVie, 94528 Rungis, France*

[d] *Unité de recherche clinique en économie de la santé (URC-ECO), hôpital Hôtel-Dieu, AP-HP, 75004 Paris, France*

[e] *Fédération francophone de cancérologie digestive (FFCD), 21000 Dijon, France*

[f] *Novartis Pharma S.A.S., 92500 Rueil-Malmaison, France*

[g] *CHU de Limoges, Département de pédiatrie, 87000 Limoges, France*

[h] *Univ. Lille, Inserm, CHU Lille, CIC 1403, Centre d'investigation clinique, 59000 Lille, France*

[i] *Département de recherche et développement clinique - Hématologie - BeiGene, Switzerland GmbH, 4051, Basel, Switzerland*

[j] *CHU Lille, Institut de pharmacie, 59000 Lille, France*

[k] *Centre d'investigation clinique 1407 INSERM - Hospices civils de Lyon, UMR 5558 CNRS université de Lyon, service de pharmacotoxicologie, 69000 Lyon, France*

[l] *Pôle USSAR, anesthésie réanimation, CHU de Poitiers, 86021 Poitiers, France*

[m] *Unité de vigilance des essais cliniques, DRCI CHRU de Nancy, 54500 Vandoeuvre lès Nancy, France*

[n] *Univ. Grenoble Alpes, Inserm, CHU de Grenoble, CIC1406, 38000 Grenoble, France*

[o] *Sorbonne Université, service de pharmacologie clinique, plateforme de recherche de l'est parisien (URCEST-CRCEST-CRB), Assistance publique - Hôpitaux de Paris (AP-HP), Hôpital St Antoine, 75000 Paris, France*

[p] *Clinsearch, 92240 Malakoff, France*

[q] *metaEvidence.org, service de pharmacotoxicologie Hospices civils de Lyon, 69000 Lyon, France*

Received 1 December 2023; accepted 4 December 2023

[*]**Corresponding author**. CHU Rennes, Inserm, Centre d'investigation clinique de Rennes (CIC1414), service de pharmacologie clinique, Institut de recherche en santé, environnement et travail (Irset), UMR S 1085, 35000 Rennes, France.
E-mail adress: clara.locher@chu-rennes.fr (C. Locher)

[1] Participants of the Round Table « Organizational and regulatory aspects » of Giens Workshops 2023.

*****Articles, analyses and proposals from Giens Workshops are those of the authors and do not prejudice the proposition of their parent organization

Summary

Clinical trials often last several months or even several years. As the trial progresses, it can be tempting to find out whether the data obtained already answers the question posed at the start of the trial in order to stop inclusions or monitoring earlier. However, knowing and taking into account interim results can sometimes compromise the integrity of the results, which is counterproductive. To minimise this risk and ensure that the treatments are assessed reliably, safety and/or efficacy criteria are monitored during the study by a Data Monitoring Committee. After receiving the results confidentially, the Data Monitoring Committee assesses the benefit/risk ratio of the study treatment

and recommends that the trial be continued, modified or terminated. Data Monitoring Committee members issuing these recommendations have an important responsibility: a hasty decision to end the trial may lead to inconclusive results unable to answer the initial question and, inversely, delaying the decision to end the trial may expose the subjects to potentially ineffective or even harmful interventions. The Data Monitoring Committee's task is therefore particularly complex. With this in mind, the round table discussion at the Giens workshops was a chance to review the scientific justification for creating Data Monitoring Committees and to recall the need for their members to receive comprehensive training on the complexities of multiple analyses, confidentiality requirements applying to the results and the need for them to be aware that recommendations to end a trial must be based on data that is robust enough to assess the benefit/risk ratio of the treatment studied.

KEYWORDS

Data Monitoring Committee; Benefit/risk ratio; Interim analyses; Integrity

Abbreviations

DMC	Data Monitoring Committee
DSMB	Data Safety and Monitoring Board
EMA	European Medicines Agency
FDA	Food and Drug Administration
GCP	Good Clinical Practice
NIH	National Institutes of Health
OS	Overall Survival
PFS	Progression Free Survival

Scientific considerations

Why monitor a clinical trial in progress?

During a clinical trial, there are different monitoring systems to protect the subjects and contribute to the quality of the trial (Fig. 1).

Firstly, the clinical trial Vigilance Unit reviews the adverse events. This monitoring, which requires expertise in pharmacovigilance or medical device vigilance, mainly involves analysing individual data to determine whether there is a causal link between the event reported and the study intervention.

For early phase trials (e.g. first-in-human trials, dose escalation studies), a committee of experts analyses the safety data to allow the inclusion of a new patient and/or proceed to the next dose level. The composition of this committee of experts is discussed in the "Early phases and DMCs" Box 1 [1-3].

Thirdly, the data accumulated during the study is analysed to check that there is still uncertainty over the advantages and disadvantages of the treatment, justifying the need to obtain further data and therefore continue the study. To answer this question, the frequency of events in the two treatment arms must be compared. However, knowing and taking into account these interim study results can sometimes compromise the integrity of the findings, which is counterproductive.

Therefore, the comparative data is evaluated by a group of experts independent from the sponsor and investigators. This group of experts is the Data Monitoring Committee (DMC). The term *Data Safety and Monitoring Board* (DSMB), which is often used in the literature, should be avoided as it focuses on safety. As explained further on, safety should be mainly assessed by looking at the benefit/risk ratio. An example is shown in Box 2 [4].

Recommendation 1

During a trial, study participants are protected by different monitoring systems, the purposes of which must be identified, and it is vital to adapt the relevant terminology. In particular, it is preferable to use the term Data Monitoring Committee (DMC) which, unlike the term Data Safety and Monitoring Board (DSMB), does not refer solely to safety data.

What are the risks inherent to interim analyses?

In order to obtain robust results, a confirmatory trial is based on hypothetico-deductive reasoning. The hypothetico-deductive approach consists of putting forward a hypothesis in advance of an *ad hoc* experiment that will test it. The trial results will confirm or disprove the initial hypothesis. A trial designed and conducted appropriately, and whose results are unbiased and interpretable, is a

trial whose initial hypothesis can be confirmed or disproved based solely on the true situation. To meet this requirement, it is vital for the tester, i.e. the investigator, to be unable to influence the results by changing the way the study is conducted.

When the benefit/risk ratio of a therapeutic intervention is evaluated during the study, i.e. before the end of inclusions or before the end of the scheduled follow-up period, it is possible to negatively influence the rest of the study and ultimately the results. When one knows the interim results it can be tempting to change the protocol, such as the rank order of efficacy endpoints or inclusion criteria. These changes based on interim results therefore impact the integrity of the study and deviate from the hypothetico-deductive approach. Knowledge of the data can also change the future treatment of study subjects, particularly in open-label studies, or affect investigators' motivation when results are temporarily inconclusive or, inversely, very much in favour of the study treatment.

Furthermore, due to random sampling fluctuations, particularly great at the start of the trial when there is little data and/or few events, the findings can suggest, temporarily, that the study treatment is effective or indeed ineffective. This is illustrated in Box 3 [5]. In this example, the final analysis shows that there is no difference between the tifacogin arm and the placebo arm. However, the retrospective analysis limited to the first 722 subjects temporarily shows that the treatment is effective. If these results are disclosed during the study, it can be tempting to stop including subjects and conclude that the treatment is effective at a moment in time when the findings do not reflect the true situation. Generally speaking, the risk of wrongly concluding that the treatment is beneficial increases the more comparisons are made. This increase in the risk of coming to a wrong conclusion when multiple tests are carried out is called overall type 1 error rate inflation. These false positive results are detrimental to patients as they can lead to the authorisation then use of treatments offering no benefits or that are no better than the standard treatment when the two are compared.

It is also important to remember that this random walk, with results temporarily supporting or discrediting the treatment's effect, affects all of the endpoints studied, including safety endpoints. Thus, results suggesting a higher risk can also be observed with a treatment that is safe, especially when a high number of adverse events are analysed (overall type 1 error rate inflation due to multiple comparisons).

To protect the scientific integrity of the study, we must therefore find a way to monitor the safety and/or efficacy data accumulated over the course of the study while applying hypothetico-deductive reasoning, and without increasing the risk of false positives through type 1 error rate inflation.

What can be done to protect the scientific integrity of the study?

To monitor safety and/or efficacy criteria during the study and at the same time preserve the study's integrity, the data must be analysed (i) confidentially, i.e. shared with a limited number of people, and (ii) by independent parties, i.e. people not involved in running the trial (investigators, steering committee, sponsor, etc.). It is with these confidentiality and independence requirements in mind that the tasks and organisation of DMCs were designed.

Indeed, a study's DMC must be composed of experts who are separate from the study, involved in no other committee (Fig. 2), and who follow operating procedures that guarantee that the production and circulation of interim results are entirely separated from the parties running the trial. This separation also allows the sponsor to propose changes to the protocol without being suspected of doing so based on the interim results. This is important because if the protocol is changed during the course of the study, it must be possible to confirm that the changes were not made to adapt to data based on interim analysis trends. DMC members must also make brief recommendations to the Steering Committee and the sponsor to stop or continue the study, ensuring that they do not reveal the result trends.

Recommendation 2

The circuit for producing and reviewing interim analysis results must remain separate from the parties running the trial. The data must therefore be analysed by an independent statistician who sends the data confidentially to the DMC members. The DMC analyses the data and recommends that the trial be continued, modified or terminated without revealing the result trends.

Is it mandatory to create a DMC?

Creating a DMC is not always necessary and is certainly not mandatory. In reality, in more cases than not, a clinical trial does not require a DMC to be created [6]. As there is no universal rule determining when a DMC is needed, together, the sponsor and the steering committee must consider the relevance of a DMC for each clinical trial. They must start considering this when

planning the study, essentially taking into account: the study population, the type of intervention and the study design. It may therefore be advisable for oversight by a DMC in the following situations: a life-threatening illness, vulnerable populations, significant or unknown risk of adverse reactions or termination rules based on the analysis of results [7]. In accordance with the decree of 2 December 2016, the "*reasons for creating or not creating a Data Monitoring Committee*" must be submitted to the ethics committee [8].

Recommendation 3

There is no universal rule to determine when a DMC is needed. The need should be considered during the design phase based on the study population, the type of intervention and the study design. If there is no DMC, this must be explained in the protocol.

Organisational aspects of DMCs

Expertise necessary to be a DMC member

Separating the circuit for producing and reviewing interim analysis results is essential but insufficient. It is true that, compared to the people involved in running the trial, the DMC members are less directly concerned with whether or not the trial's findings are positive. That said, the random walk of results can also lead to DMC members making inappropriate decisions concerning efficacy or safety, and, without sufficient training, they may jeopardise the trial's scientific integrity.

To avoid this pitfall, the DMC must avoid making decisions naively or intuitively. They must follow a rigorous process based on appropriate comparative statistical methods. For example, to avoid overall type 1 error rate inflation caused by repeat testing of efficacy criteria, methods are needed to adjust the significance threshold (e.g. Lans and DeMets, Peto, etc.). Concerning safety, no decision to terminate the trial should be made simply based on a list of adverse events, or by merely observing a large difference even if it is nominally significant (*P* value <0.05) or if it concerns mortality (which is just as subject to random sampling fluctuations as other criteria). The problem of type 1 error rate inflation with multiple safety criteria persists and makes the DMC's task complicated.

The DMC must therefore be aware of the concept of random sampling fluctuations and avoid making hasty decisions, which would make the trial results inconclusive and unable to answer the initial question, but also avoid delaying the decision too much, which may expose subjects to potentially ineffective or even harmful interventions [9]. It is therefore vital to create a specific training program. Like "good clinical practice [GCP]" training courses validating knowledge, this training could be recommended or even made compulsory for DMC members.

Recommendation 4

DMC members must be made aware of the requirement to keep the results confidential, of the issues of multiple analyses and of the responsibilities they take when issuing their recommendations. It is therefore vital to create a specific training program.

Composition of the DMC

When it is confirmed that a DMC is needed, the ideal number of members is three or five. When there are more than five members, planning problems can prevent the committee from functioning properly. The meeting schedule takes into account the number of topics to include, the interim analyses scheduled and study deadlines. Members must commit for the entire duration of the study. Depending on the needs of the study, the experts are generally clinicians, plus at least one biostatistician or methodologist or expert in clinical development and data analysis. The DMC members must have experience in the medical field studied, including clinical research, and be independent from the study. An independent expert means that they have nothing to lose or gain in recommending that the study be ended or continued, including on a personal level if they are involved in a similar trial. Ideally, it is therefore preferable to avoid including members involved in competing trials.

Recommendation 5

The DMC members must not be involved in conducting the study and must be separate from the sponsor and have no interest in the study's continuation or termination.

Data needed to issue recommendations

At the time of the meeting, the DMC must have all of the (ideally monitored) data to make a decision. The sponsor must take into account the DMC meeting schedule to anticipate what must be done to produce reliable data.

The data must be presented aggregated by treatment arm, preserving the anonymity of the arms as much as the investigation procedure allows. If the data is incomplete, the data's presentation must make it possible to assess how representative the data is.

The reliability of the data depends both on the quality of the data collected and on the analysis conducted, which must leave no doubt over the interpretation and must be presented clearly in the form of a report organised in relation to the main and secondary objectives.

Recommendation 6

The DMC must have reliable data to be able to issue reliable recommendations. The sponsor must therefore (i) make every effort to ensure that all the data needed by the DMC is up to date and of a high quality and (ii) comply with the monitoring schedule.

Tasks potentially entrusted to the DMC

A DMC can be created to issue recommendations to stop or continue the study due to safety data, efficacy data and/or futility, depending on the tasks it is given in the protocol and in a charter created before the study commences.

Recommendations to stop the trial due to a lack of safety

The safety analysis should be designed to monitor the benefit/risk ratio. Indeed, a higher risk of serious adverse reactions can be considered acceptable if the new treatment's efficacy offers patients a significant clinical benefit. This is the case, for example, for new anticoagulants which

provide a greater benefit than previous generations but also cause more major bleeding (Box 4) [10,11]. In similar situations, monitoring safety without considering the benefits would lead to the trial being prematurely terminated due to a lack of safety, which would deprive patients of treatments with a greater net clinical benefit than existing treatments.

This assessment requires being able to compare the safety and efficacy for events of the same clinical seriousness. To do this, a primary safety endpoint is chosen with the same clinical seriousness as the primary efficacy endpoints. In other cases, a net clinical benefit criterion is used as the primary endpoint. All-cause mortality also makes it possible to assess the net clinical benefit when adverse reactions are potentially fatal.

One difficulty to take into account when assessing the benefit/risk ratio is the frequent time lag between (i) the onset of adverse events, which can occur soon after initiating treatment, whether this is surgery, an implantable device or a drug, and (ii) the point at which a clinical benefit is observed. For example, for a first-line cancer treatment, adverse reactions may occur when the treatment is administered at the start of the trial. But several months or even several years of follow-up are needed to demonstrate the benefit in terms of survival, as death generally occurs after several progressions and several lines of treatment.

In addition to these time lag difficulties, as a precaution, a suspected safety issue is considered a lack of safety without requiring formal proof. But it is vital to bear in mind that safety monitoring is complicated by the risk of "false positives" due to (i) random sampling fluctuations and (ii) multiple comparisons. There are considerable random sampling fluctuations particularly when there are few subjects, i.e. at the start of the trial, or when few events are observed (Box 5) [12,13]. The "false differences" that arise from these fluctuations can cause the DMC to wrongly conclude that there is a higher risk of events. As for multiple comparisons, this is due to multiple types of adverse events or even different groups of events being compared, which will also increase the risk of "false positives".

Recommendation 7

When recommending that a trial be ended for safety reasons, the DMC must ensure that the trial data is sufficiently robust to conclude that the benefit/risk ratio is negative.

Recommendations to end a trial for efficacy reasons

At first glance, efficacy monitoring may seem less complicated and less subjective than safety monitoring because it is based on statistical rules provided in the protocol (Lan and DeMets, Peto-Haybitte, etc.) aimed at controlling type 1 error rate inflation. As a reminder, these termination rules must be established beforehand (Box 6) [14].

Initially, interim efficacy analyses were used to stop a trial early on, before all patients were included, as soon as efficacy was demonstrated for the primary endpoint. But, it is important to remember that stopping a trial too early is likely to compromise the interpretation of the benefit/risk ratio. Therefore, even when efficacy is demonstrated for the primary endpoint, it may be necessary to continue including subjects if the benefit/risk ratio is uncertain (Box 7, Fig 3.) [15].

With an appropriate overall type 1 error control plan, efficacy can now be demonstrated based not on just one but several endpoints. The "termination rules" can then be used in a different way: they make it possible to demonstrate one or more endpoints early on without ending the trial. The trial can then be continued to provide data to reach a conclusion regarding all of the endpoints. The demonstrated efficacy for the first endpoint(s) may be used to submit a marketing authorisation application to the competent authorities. For example, endpoints frequently used in oncology are: (i) progression-free survival (PFS) and (ii) overall survival (OS). In pivotal studies, these two criteria are often ranked in hierarchical order, PFS first then OS. The challenge is obviously to show a benefit in terms of OS but data maturity takes more time than for PFS (Box 8) [16].

When interim analyses are foreseen, it is also important not to confuse the new treatment's efficacy for a safety problem with the control treatment. Take, for example, a new treatment that reduces mortality compared to the reference treatment. At the start of the trial, there is not enough information to conclude that there is a benefit in terms of mortality: the termination rule is not met. However, at this stage, a higher number of deaths in the control group could suggest that the reference treatment causes a higher mortality rate, which could lead to the study being terminated as a precaution due to a safety problem in the control group. But since this is the reference treatment, subjects randomised to the control group would have received the same treatment if they had not taken part in the study. Continuing the study therefore does not constitute a loss of chance for subjects who will be randomised to the control group. But misinterpreting the situation and stopping the trial as a precaution (i) means not being able to demonstrate the benefit of the new treatment in

terms of mortality and (ii) poses an ethical problem, because the subjects are exposed to the constraints of a study whose results will not provide satisfactory scientific and medical insights. Other studies will be needed to provide the answer to the question posed, which will delay the progress of knowledge.

In conclusion, there is no urgent need to stop a trial for efficacy, as tempting as this may be. If the trial has started, this is because everything planned (the number of subjects to be included and the duration of follow-up) was considered ethically acceptable and necessary given that the new treatment shows sufficient promise of potential therapeutic progress. It would therefore be detrimental to end it prematurely before it provides evidence robust enough to change practices and use the new treatment. Ending the trial in this way also poses two ethical problems: exposing subjects to the potential risks of the new treatment for nothing (Box 9) [17] and depriving subjects of therapeutic progress due to lack of proof as the study's scientific integrity has been compromised.

Stopping a trial prematurely for efficacy to give patients quicker access to the new product is highly questionable. This means potentially interrupting the patients' treatment, as continuing treatment will require exceptional access that can be tricky to set up. In addition, access to the new product will not be based on a single study but on all of the safety and efficacy data from studies carried out in the context of the new product's development plan. Evaluating the benefit/risk ratio of the new treatment based on all of the available data takes time, after which the relevance of national welfare scheme coverage must be assessed then the reimbursement and pricing must be set. Stopping the trial too early can actually be counterproductive for patients, weakening the proof of safety and efficacy and therefore jeopardising access to the new product due to an insufficient level of proof. Some early terminations for efficiency are actually driven by financial and competitive interests rather than the interests of patients [18].

Stopping a trial too early for efficacy can invalidate the results as the following problems lead to weak and unrobust results:

- Shorter exposure, limiting the safety assessment;
- An imprecise estimation of the treatment's effects (wide adjusted confidence interval affecting the absence of effect);
- An overestimated treatment effect [19];
- An unrealistic effect size or analysis performed with an unrealistic fraction of information;
- Difficulty assessing the consistency of results between endpoints and subgroups;

- Imprecise Kaplan-Meier curves with censoring linked to a very early cutoff date;
- etc.

Recommendation 8

When recommending that a trial be ended for efficacy reasons, the DMC must ensure that the trial data is robust enough to make it the standard of care.

Recommendations to stop the trial due to futility

When the trial evaluates a medical practice that has become common without any proven benefit, the trial should not be ended for futility reasons (the protocol should not even foresee this). If this treatment does not provide any benefit, it must be possible to provide a level of proof convincing enough to end a medical practice already established. Saying that the trial failed to demonstrate the intervention's benefit will not have the same significance as (i) a result showing a lack of effect with a narrow confidence interval or (ii) a deleterious effect. In this situation the trial must be completed, even if safety considerations arise. When there is already an established medical practice, the community will not change practices if the trial does not provide strong proof of a lack of safety (Box 10) [20,21]. A poorly established higher risk is not necessarily enough to convince the community. Therefore, even if one is less effective or less safe than the other, participating in the study does not result in a loss of chance for these patients compared to the usual care.

Recommendation 9

When recommending that a trial be ended for futility reasons, and when there is already an established medical practice, the DMC must ensure that the trial data is robust enough to decide that the intervention has no benefits.

The DMC's recommendations must also take the timeline into account. The recommendation may need to be issued urgently when a study must be terminated based on safety

data, for the good of the subjects included, while a recommendation for efficacy or futility is less urgent.

Publication of interim analysis results

From a methodology point of view, it is universally acknowledged that the results of inconclusive interim analyses should not be disclosed as they may compromise the study's scientific integrity. However, these interim analysis results are quite often published, particularly in oncology. As described above, the two main endpoints used in oncology are progression-free survival and overall survival. There is a time lag between when the benefit is demonstrated for these two endpoints because progression-free survival events include radiological progression, which occurs more quickly than death. Interim analysis plans take this time lag into account and a benefit in terms of progression-free survival is demonstrated before a benefit in terms of overall survival. Given that demonstrating a benefit in terms of progression-free survival offers the chance to seek authorisation from the authorities for this new drug, these results are regularly shared during conferences and published.

Publishing progression-free survival results does not raise methodological problems because they are definitive. However, these publications also typically report inconclusive overall survival results from the interim analysis. This is because overall survival results are usually analysed at the same time as progression-free survival. Because of this, papers specify that one should not take into account the negativity of this OS analysis as the data is not mature. In other cases, the findings may be misrepresented if the effect on overall survival is nominally significant (Box 11) [22]. Although inconclusive, the findings are represented in a way to suggest that overall survival has already been demonstrated.

Sharing inconclusive results in this way, suggesting a potentially positive trend, can affect the conduct of the trial and compromise its integrity in relation to its main objective, which is to demonstrate a benefit in terms of overall survival. Investigators, and indeed the patients, may not realise the uncertainty of this trend, and investigators may want to take the patients out of the study to prescribe the study treatment if it is already available. These premature study withdrawals mean potentially not being able to reach a conclusion about this endpoint and lingering uncertainty over the treatment's benefit in terms of overall mortality.

Recommendation 10

When the statistical analysis plan foresees a hierarchical analysis of the endpoints, only conclusive analysis findings should be disclosed. The inconclusive endpoint outcomes for which the study is continued must not be disclosed.

Importance of the DMC's work

DMC members help to keep clinical trial subjects safe and protect the scientific integrity of this research. Given their scientific and medical responsibilities, expertise and time invested, the work of DMC members is vital and an activity of general interest that should be recognised and valued. DMC members are often health professionals working in hospitals, universities and university hospitals and have a heavy workload. Their work for DMCs could be recognised through career advances (promotions, bonuses):

- Professional recognition: for hospital and academic staff, by allocating time to clinical research.
- Financial compensation: possible.
- Publication: DMC members must be informed that they will not be listed as authors of the publication.

Recommendation 11

The time and expertise invested for DMC tasks (unpaid institute-sponsored trials) should be valued as collective activities of general interest (as is the case when participating as a member or president of an ethics committee, or as an expert assessor of health products for the ANSM, EMA, HAS, etc.) and should be taken into account when they apply for a hospital or university promotion.

In conclusion, DMCs help to protect people participating in a trial. However, not all clinical trials need a DMC. The decision whether or not to use a DMC must therefore be anticipated and justified

from the study design phase. If the trial requires a DMC, the data must be monitored with the greatest methodological care to avoid compromising the study's scientific integrity and the significance of the results by jeopardising their credibility. Before the meetings, the sponsor must make every effort to ensure that the data needed to issue recommendations is up to date and of a high quality. For their part, DMC members must be aware of their responsibility and the complexity of the task entrusted to them. Their decisions will need to take into account methodological issues inherent to the hypothetico-deductive approach but also the statistical aspects of repeat analyses. These decisions require DMC members to have solid expertise and to follow all of the logistical organisation rules for interim analyses. This is why DMC members must have received comprehensive training on scientific issues that they must take into account when devising their recommendations.

Disclosure of interests

Authors have no competing of interest to declare

References

[1] EMA. Data monitoring committees - Scientific guideline. European Medicines Agency 2018. https://www.ema.europa.eu/en/data-monitoring-committees-scientific-guideline. [Accessed October 11, 2023].

[2] Adashek JJ, LoRusso PM, Hong DS, Kurzrock R. Phase I trials as valid therapeutic options for patients with cancer. Nat Rev Clin Oncol 2019;16:773–8.

[3] Van Norman GA. Data safety and monitoring boards should be required for both early- and late-phase clinical trials. JACC Basic Transl Sci 2021;6:887–96.

[4] Alexander JH, Lopes RD, James S, Kilaru R, He Y, Mohan P, et al. Apixaban with antiplatelet therapy after acute coronary syndrome. N Engl J Med 2011;365:699–708.

[5] Abraham E, Reinhart K, Opal S, Demeyer I, Doig C, Rodriguez AL, et al. Efficacy and safety of tifacogin (recombinant tissue factor pathway inhibitor) in severe sepsis: a randomized controlled trial. JAMA 2003;290:238–47.

[6] EMA. Data monitoring committees - Scientific guideline. European Medicines Agency 2018. https://www.ema.europa.eu/en/data-monitoring-committees-scientific-guideline. [Accessed October 11, 2023].

[7] Evans SR. Independent oversight of clinical trials through data and safety monitoring boards. NEJM Evidence 2022;1:EVIDctw2100005. https://evidence.nejm.org/doi/full/10.1056/EVIDctw2100005. [Accessed October 11, 2023].

[8] Legifrance. Arrêté du 2 décembre 2016 fixant le contenu, le format et les modalités de présentation du dossier de demande d'avis au comité de protection des personnes sur un projet de recherche mentionnée au 1° de l'article L. 1121-1 du code de la santé publique portant sur un médicament à usage humain. December 2016. https://www.legifrance.gouv.fr/loda/id/JORFTEXT000033545400. [Accessed October 11, 2023].

[9] Schulman S, Carlsson A, Laporte S. When is it safe to stop for safety? J Thromb Haemost. 2012;10:1188-90.

[10] Wiviott SD, Braunwald E, McCabe CH, Montalescot G, Ruzyllo W, Gottlieb S, et al. Prasugrel versus clopidogrel in patients with acute coronary syndromes. N Engl J Med 2007;357:2001–15.

[11] ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction: ISIS-2. Lancet 1988;332:349–60.

[12] Food and Drug Administration. Research C for DE and. FDA Drug Safety Podcast: Interim clinical trial results find increased risk of leg and foot amputations, mostly affecting the toes, with

the diabetes medicine canagliflozin (Invokana, Invokamet); FDA to investigate. 2022. https://www.fda.gov/drugs/fda-drug-safety-podcasts/fda-drug-safety-podcast-interim-clinical-trial-results-find-increased-risk-leg-and-foot-amputations. [Accessed October 11, 2023].

[13] Arnott C, Huang Y, Neuen BL, Di Tanna GL, Cannon CP, Oh R, et al. The effect of canagliflozin on amputation risk in the CANVAS program and the CREDENCE trial. Diabetes Obes Metab 2020;22:1753–66.

[14] De Bruyne B, Pijls NHJ, Kalesan B, Barbato E, Tonino PAL, Piroth Z, et al. Fractional flow reserve–guided PCI versus medical therapy in stable coronary disease. N Engl J Med 2012;367:991–1001.

[15] Eikelboom JW, Connolly SJ, Bosch J, Dagenais GR, Hart RG, Shestakovska O, et al. Rivaroxaban with or without aspirin in stable cardiovascular disease. N Engl J Med 2017;377:1319–30.

[16] Mirza MR, Chase DM, Slomovitz BM, dePont Christensen R, Novák Z, Black D, et al. Dostarlimab for primary advanced or recurrent endometrial cancer. N Engl J Med 2023;388:2145–58.

[17] Nissen SE, Wolski KE, Prcela L, Wadden T, Buse JB, Bakris G, et al. Effect of naltrexone-bupropion on major adverse cardiovascular events in overweight and obese patients with cardiovascular risk factors: a randomized clinical trial. JAMA 2016;315:990–1004.

[18] Iltis AS. Stopping trials early for commercial reasons: the risk–benefit relationship as a moral compass. J Med Ethics 2005;31:410–4.

[19] Bassler D, Briel M, Montori VM, Lane M, Glasziou P, Zhou Q, et al. Stopping randomized trials early for benefit and estimation of treatment effects: systematic review and meta-regression analysis. JAMA 2010;303:1180–7.

[20] Combes A, Hajage D, Capellier G, Demoule A, Lavoué S, Guervilly C, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. N Engl J Med 2018;378:1965–75.

[21] Harrington D, Drazen JM. Learning from a trial stopped by a data and safety monitoring board. N Engl J Med 2018;378:2031–2.

[22] Benboubker L, Dimopoulos MA, Dispenzieri A, Catalano J, Belch AR, Cavo M, et al. Lenalidomide and dexamethasone in transplant-ineligible patients with myeloma. N Engl J Med 2014;371:906–17.

Legends of figures

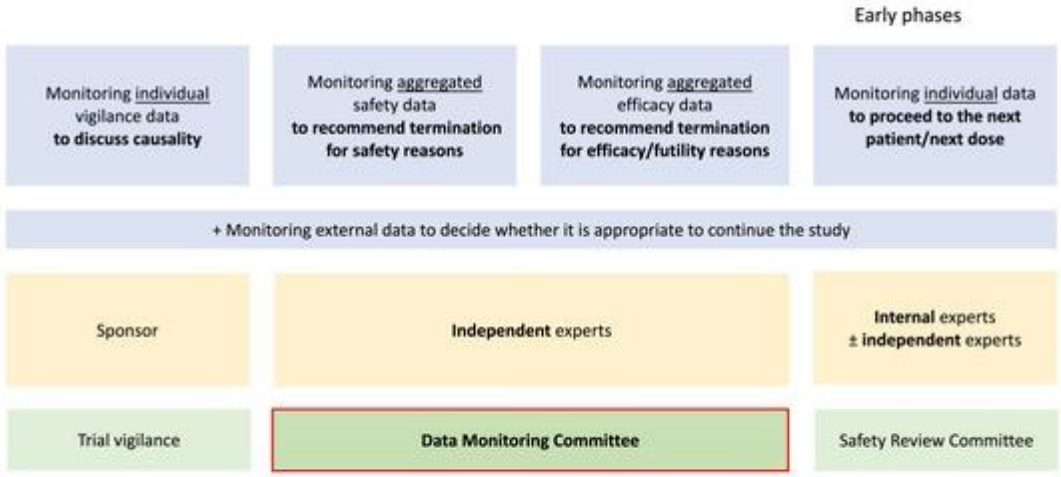

Figure 1. Monitoring systems in place to protect study subjects.

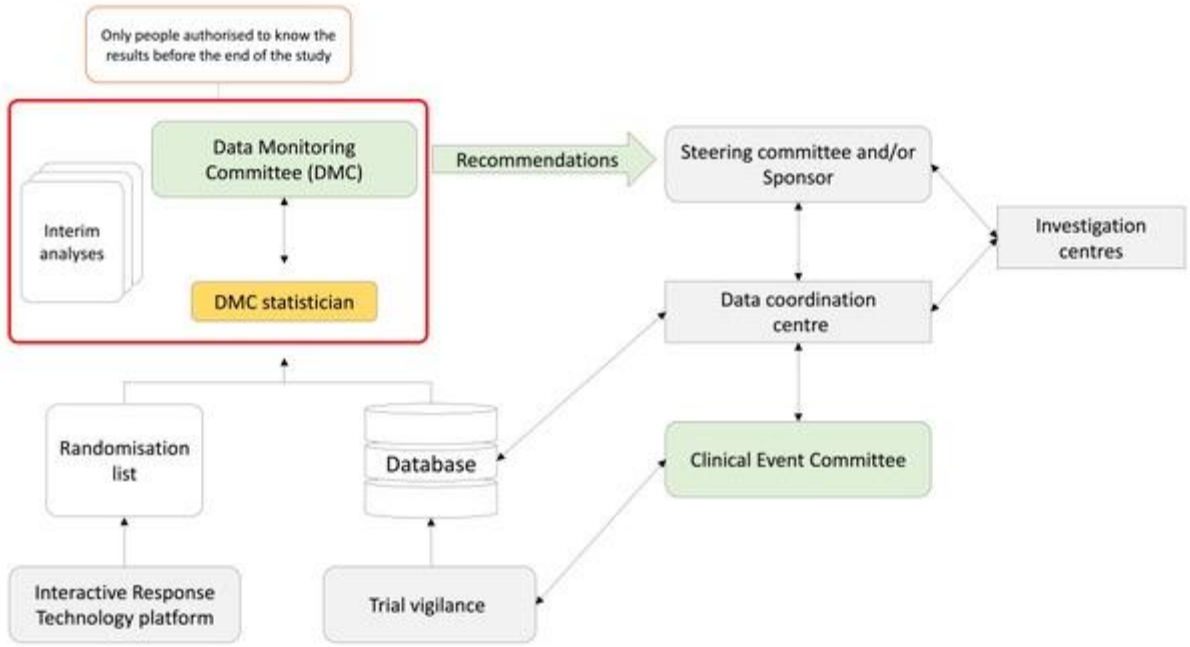

Figure 2. Committees and people involved in a trial. The DMC members must be independent and as such, they must not be involved in any committee other than the DMC.

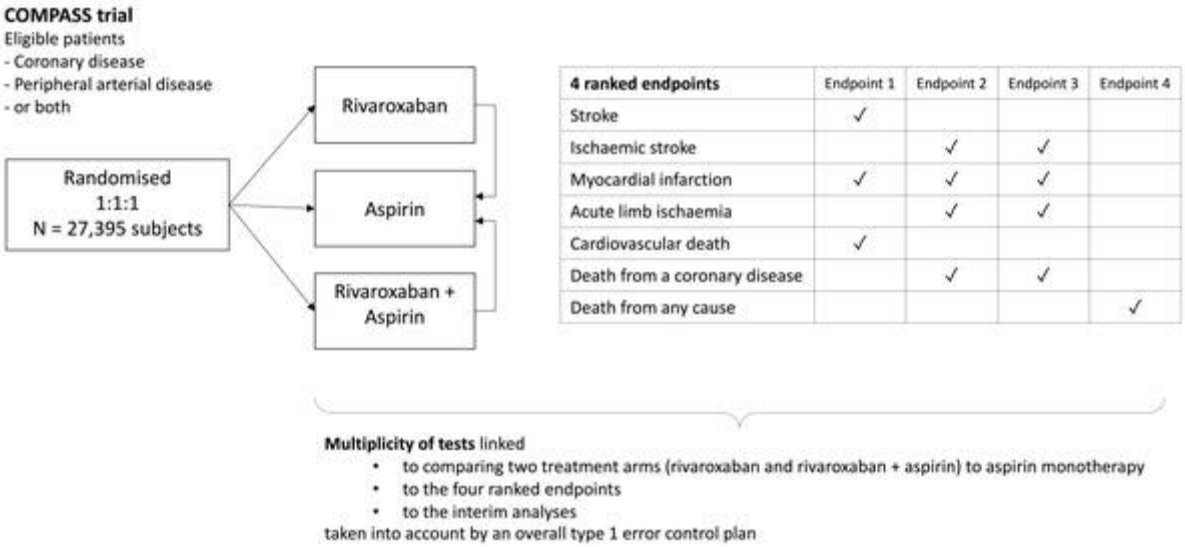

Figure 3. COMPASS trial design. The analysis strategy planned to test four ranked endpoints by comparing the two methods of use of rivaroxaban (monotherapy or combined with aspirin) against aspirin monotherapy.

Box 1. Early phases and Data Monitoring Committees (DMCs)

The Food and Drug Agency (FDA) and European Medicines Agency (EMA) strongly recommend creating a DMC for confirmatory (phase III) trials. Since the late 1990s, NIH has even required a DMC for all confirmatory trials that it sponsors. However, the agencies' recommendation on the need for a DMC for early phase trials is less clear.

Phase I trials, which are to assess safety, determine the dose and evaluate the pharmacokinetic properties of the drug, rely on monitoring committees responsible for reviewing events that may be dose-limiting toxicities, and making decisions about dose escalation. Unlike confirmatory trial DMCs, these committees, often called *Safety Review Committees*, may have members who are not independent from the sponsor and the investigator, and the steering committees for these early trials often include external members [1]. For phase II trials, the goal of which is to obtain preliminary efficacy data, it is also unusual to have a DMC.

However, the boundaries between the different clinical development phases of a drug candidate are becoming increasingly blurred, and nowadays trials are often designed to meet multiple objectives. This is the case for combined phase I/II trials, which often combine a dose escalation phase with one or more extension cohorts, which may or may not be randomised, aimed at providing preliminary efficacy data. In some cases, particularly in oncology, these trials may be pivotal studies [2]. Whether the trials are combined phase I/II trials or adaptive phase II trials, interim analyses can provide efficacy results at a very early stage of development. However, the variability of these early results means a higher risk of coming to the premature and incorrect conclusion that the treatment is beneficial. Without external expertise from parties outside of the trial steering committee and the sponsor, there is a risk of prematurely ending a trial that would have provided negative results had it been continued. For early phase trials, this risk is acceptable when the marketing authorisation is subject to a phase III trial confirming the benefit. But when early phase trials act as pivotal studies, the creation of a DMC is fully justified [3].

Box 2. The APPRAISE-2 trial [4] comparing apixaban to placebo in addition to antiplatelet treatment in acute coronary syndromes was terminated prematurely by the DMC after including around 7,000 subjects. The reason for this early termination was an unfavourable benefit/risk ratio due to an increase in major bleeding (HR = 2.59 [1.50 – 4.46]) not counterbalanced by a decrease in ischaemic events (HR = 0.95 [0.80 – 1.11]).

Box 3. The OPTIMIST trial [5] assessed the 28-day mortality rate of tifacogin, a human recombinant tissue factor pathway inhibitor, in severe sepsis.
Retrospective analysis limited to the **first 722 subjects included:**
- 38.9% in the placebo arm
- 29.1% in the tifacogin arm (*P* = 0.006)

Analysis of the **1,754 subjects included**:
- 33.9% in the placebo arm
- 34.2% in the tifacogin arm (*P* = 0.88)

Stopping the trial after including the first 722 subjects would have led to the incorrect conclusion that the treatment was beneficial.

Box 4. The TRITON–TIMI 38 trial [10] evaluated prasugrel, a new antiplatelet drug, compared to clopidogrel in acute coronary syndrome. A 32% relative increase in major bleeding was observed (HR = 1.32 [1.03 – 1.68]), although with a 19% relative decrease in ischaemic events (HR = 0.81 [0.73 – 0.90]).
The ISIS-2 trial [11] evaluated streptokinase and/or aspirin in the acute phase of coronary syndrome using a factorial design. Seven cerebral haemorrhages were observed in the streptokinase group versus none in the placebo group, but this greater number of cerebral haemorrhages was accompanied by a significant decrease in vascular mortality at 5 weeks (OR = 0.75 [0.68 – 0.83]).

Box 5. The CANVAS trial compared canagliflozin to a placebo in type 2 diabetes. The primary endpoint was non-inferiority in cardiovascular events. During an interim analysis, the DMC identified a higher number of amputations in the canagliflozin groups: 7 out of 1,000 patients in the canagliflozin 100mg/day group; 5 out of 1,000 patients in the canagliflozin 300mg/day group; 3 out of 1,000 patients in the placebo group [12]. This risk was not observed in another similar trial – the CANVAS-R trial – also currently in progress. The DMC decided to continue the trial, but the agencies added amputations to the list of adverse effects for canagliflozin. After this interim analysis, the safety data from the CANVAS and CREDENCE trials, in diabetic nephropathy, were grouped together and after an in-depth analysis of the data, it was concluded that the higher risk of amputation initially observed in the CANVAS trial was probably random [13].

Box 6. The FAME 2 trial [14] evaluated fractional flow reserve-guided percutaneous coronary intervention in patients with stable angina. The abstract states that: "*Recruitment was halted*

prematurely after enrollment of 1,220 patients (888 who underwent randomization and 332 enrolled in the registry) because of a significant between-group difference in the percentage of patients who had a primary end-point event". And yet the protocol specified ("*Interim analysis: no interim analysis is planned*") (page 26 section 13.3.2).

The reason given for premature termination is a significant between-group difference. But the principal investigator should never have known about this difference given that there was no process for handling the multiplicity generated by potential interim analyses. This decision to end the trial was therefore based on the results and it is impossible to rule out the possibility that it was merely a particularly favourable random fluctuation not reflecting the true efficacy of the treatment.

Box 7. The COMPASS trial [15], the design of which is presented in **Figure 3**, compared the benefit of rivaroxaban or its combination at a reduced dose with aspirin in relation to aspirin monotherapy in secondary cardiovascular prevention. The overall type 1 error control plan considered the possibility of providing demonstrations for several endpoints including all-cause mortality.

During the first interim analysis, the DMC recommended stopping the study for efficacy reasons due to a difference for the first endpoint in the rank order. This difference in cardiovascular events was observed only in the rivaroxaban + aspirin arm and not in the aspirin monotherapy arm (HR = 0.76 [0.66 – 0.86]; *P* <0.001). Despite no difference between the rivaroxaban monotherapy arm and the aspirin monotherapy arm (HR = 0.90 [0.78 – 1.03]; *P* = 0.12), inclusions in the rivaroxaban arm were also stopped.

As planned, the fact that the difference between rivaroxaban and aspirin was not significant for the first endpoint meant that the analysis was stopped and therefore the following endpoints could not be tested. "*The study protocol and Statistical Analysis Plan do not specify modifications to the Hochberg gatekeeping procedure for the testing of secondary efficacy outcomes in the case of a premature termination for efficacy when both comparisons are stopped but only one of the comparisons met the modified Haybittle-Peto boundary*".

Overall, the only demonstrated result, the reduction in fatal and non-fatal cardiovascular events, does not make it possible to conclude that there is a favourable benefit/risk ratio. This is because this effect on cardiovascular events is completely counterbalanced by an increase in major bleeding, which was the primary safety endpoint (12 more major bleeding events for 13 fewer ischaemic events per 1,000 subjects treated for 23 months).

This higher number of major bleeding events may have been acceptable if a decrease in all-cause mortality (endpoint ranked fourth in the hierarchy) had been demonstrated. But the decision to stop the COMPASS trial early based on the DMC's recommendations makes it impossible to analyse this endpoint. And while a trend appears to be emerging (HR = 0.82 [0.71 – 0.96]), it should not be

considered statistically significant and should not be factored into the decision. What's more, the p-values of the secondary endpoints should not appear in the table of results as, according to the statistical analysis plan, these analyses should not have been carried out.
This example perfectly illustrates that a study must only be stopped prematurely when the results robustly and irrefutably show a favourable benefit/risk ratio (and not just a certain level of efficacy).

Box 8. The RUBY trial [16] is evaluating dostarlimab in combination with platinum-based chemotherapy for the treatment of adult patients with advanced or recurrent endometrial cancer. After a median follow-up of 25 months, the interim analysis results for all subjects included show an effect on PFS (HR = 0.64 [0.51 – 0.80]). However, the OS outcomes are not significant and follow-up is being continued to determine whether there is a positive effect on mortality. The RUBY trial is therefore still in progress but given the demonstrated benefit in terms of PFS, dostarlimab has already been granted early access authorisation in this indication.

Box 9. The LIGHT trial [17] evaluated the occurrence of cardiovascular events in overweight or obese subjects treated with naltrexone + bupropion versus placebo. To avoid an increased risk of cardiovascular events, the FDA requested a non-inferiority trial with a non-inferiority margin, corresponding to the upper limit of the HR confidence interval, of 1.4. An interim analysis was also foreseen after 25% of expected events had occurred and the study would be continued only if the upper limit of the HR confidence interval was less than 2.
But because the sponsor published this interim analysis, breaching confidentiality, it was decided to end the study without being able to assess non-inferiority. Hence, despite the inclusion of 8,910 subjects, the LIGHT trial does not answer the question posed at the start of the trial and it is impossible to rule out the possibility that the naltrexone + bupropion combination causes a higher risk of cardiovascular events.

Box 10. The trial by Combes et al. [20] compared the 60-day mortality of patients with SARS treated with ECMO versus standard ventilation. The DMC stopped the trial for futility reasons when 73% of the information was obtained because the futility threshold set in the protocol had been reached. The mortality outcome was a relative risk of 0.76 [0.55 – 1.04]; *P* = 0.09. This outcome, bordering on demonstrating a mortality benefit, is difficult to interpret, and although the rules were applied, it is unfortunate that the study was stopped for futility reasons. Clearly the results do not show a clear lack of benefit with ECMO, but nor does it allow one to conclude that it is beneficial. This example is a reminder that statistical rules should only be used to guide the DMC and that it needs to look at the big picture [21].

Box 11. The FIRST trial [22] illustrates this type of misrepresentation: "*Although the difference in overall survival did not cross the prespecified superiority boundary (P<0.0096), continuous lenalidomide–dexamethasone reduced the risk of death, as compared with MPT (hazard ratio, 0.78; 95% CI, 0.64 to 0.96; P = 0.02)*".