A response by Servier to the Statement of Reasons provided by NICE

Servier gratefully acknowledges the Statement of Reasons document from NICE, and is pleased to provide information to assist NICE in its appraisal of strontium ranelate. This document summarises key information relating to the queries stated by NICE, and provides a detailed response to each request for information. In addition, a full description of the correspondence with the EMA is provided in Appendix A.

Summary of Information Required

Point	Issue	Response	Full Response See Page
4	Were there multiple subgroup analyses of TROPOS	Only 1 subgroup analysis on first request by EMA	5 main document Appendix 1
5	All documents to / from EMA on subgroup	Provided	5 main document Appendix 1
7	Justification for age cut off	Provided	7 main document
8	Biologic basis for greater benefit	Provided	8 main document Appendix 2
9	All sub group analyses performed	Provided	9 main document Appendix 1
13	Rationale for use of whole population and sub group data	Provided	11 main document
14	Statistical rationale use of whole population and sub group data	Provided	11 main document
15	Statistical analyses of TROPOS population after sub group removed	Explanation as to why this is not provided	12 main document

PROTELOS TIMELINE

1990

Donaldson LJ *et. al.* reported that the incidence of hip fracture rises exponentially after 74 yr

1996

The phase III program was designed and set up – TROPOS

With knowledge of the draft guideline on peripheral fracture recommendations

1st patient enrolled November 1996

1997

CPMP guideline on primary osteoporosis

"Peripheral fracture was the recommended efficacy end-point."

2001

CHMP guideline on primary osteoporosis (CPMP/EW/552/95 rev 1)

"Hip fracture included in the guidelines"

TROPOS study underway

25 June 2003

Protelos marketing authorization submission to EMA

Day 120 questions (November 2003)

In response to a request made by the EMA in the Day 120 questions, *a single post hoc analysis* was carried out on a subgroup of women at elevated risk of hip fracture to establish efficacy of Strontium ranelate for hip fracture

Day 180 questions (April 2004)

EMA requested specific additional supportive analyses to be performed to further explore the efficacy of strontium ranelate for the prevention of hip fractures

21 September 2004

The European Commission issued a Marketing Authorisation for PROTELOS

Position statement

In a pivotal Phase III study, TROPOS, strontium ranelate was demonstrated to significantly reduce the incidence of osteoporosis-related peripheral fractures, compared to placebo. Strontium ranelate therefore represents a valuable treatment option for post-menopausal women at risk of osteoporotic fracture.

Since TROPOS was conducted, there has been increasing emphasis on hip fracture as a key measure of efficacy for osteoporosis. Consequently, and at the request of the EMA, a *post hoc* subgroup analysis was performed to evaluate the efficacy of strontium ranelate in preventing hip fractures. This analysis found that strontium ranelate gave rise to a significant reduction in the risk of hip fracture, compared to placebo, for women aged ≥74 years with osteoporosis.

When conducting *post hoc* analyses it is vital to ensure that the analysis is robust and statistically sound. As the incidence of hip fracture in the study population was low, evaluation of hip fractures across the whole study population had insufficient statistical power to provide valid and meaningful results. Therefore, a subgroup was identified in which the risk of fracture was sufficiently high as to facilitate a robust and fully powered analysis. The subgroup was identified by screening the placebo arm of the study for the effects of known risk factors on fracture incidence. This approach ensured that the subgroup selection was not influenced by the efficacy of strontium ranelate, and allowed a single analysis to be performed, without the need for multiple exploratory analyses.

The selection criteria for the analysed subgroup were both biologically and statistically justified:

Age ≥74 years
- Age is a known risk factor for osteoporosis, and the risk of fracture increases exponentially after the age of 74 years.[1]
- The age cut-off of 74 years was selected as it is consistent with the inclusion criteria of the study, and is in line with established evidence on fracture risk.
- In the placebo arm of TROPOS, women aged ≥74 years had a significantly higher risk of hip fracture than younger women.

BMD T-score ≤-3

- BMD T-score is significantly associated with fracture risk.[2]
- The selected BMD score cut-off value of ≤-3, based on the local normative data, is equivalent to a T-score of ≤-2.4 using the NHANES III normative data. NHANES data was used as the reference values in the major osteoporosis outcome trials for bisphosphonates and in TA160/161[3,4]. Hence a T-score of ≤-3 in the TROPOS study is closely aligned with the internationally accepted criteria for osteoporosis (NHANES III ≤-2.5)
- In the placebo arm of TROPOS, a T-score ≤-3 was associated with a significantly higher risk of hip fracture than a T-score >-3.

The *post hoc* subgroup analysis of strontium ranelate in women aged ≥74 years with osteoporosis therefore provides a valuable and robust assessment of the efficacy of strontium ranelate in preventing hip fractures.

It is acknowledged that extrapolation from such a subgroup to a broader population is challenging. It is therefore understood that NICE may wish to limit their recommendations to a corresponding population.

The findings from TROPOS as a whole and the *post hoc* subgroup analysis provide a comprehensive evaluation of the efficacy of strontium ranelate for the treatment of osteoporosis. These two analyses provide different and complementary information, and Servier encourage NICE to consider both evaluations when assessing the respective aspects of osteoporosis treatment.

Areas

Subgroup analyses

The clinical efficacy of strontium ranelate has been demonstrated in two double-blind, randomised, placebo-controlled studies. The Spinal Osteoporosis Therapeutic Intervention (SOTI) study and the Treatment Of Peripheral Osteoporosis Study (TROPOS) investigated the efficacy of strontium ranelate for the prevention of osteoporosis-related vertebral and non-vertebral fractures, respectively. Both of these trials were designed and initiated in 1996, when the CPMP guideline for primary osteoporosis (CPMP/EWP/552/95) was at draft stage. Both trials were designed in line with these guidelines, which recommended that the prevention of peripheral, non-axial fractures should be used as a key measure of efficacy. Therefore, TROPOS was designed with the incidence of all peripheral fractures as the primary focus.

The inclusion criteria for TROPOS were:
- ambulatory Caucasian postmenopausal women
- aged ≥74 years
- aged 70–74 years with at least one additional clinical risk factor
- femoral neck bone mineral density (BMD) T-score <-2.5.[*]

In subsequent revisions of the CPMP guidance (several years after enrolment for TROPOS began) the recommended primary efficacy endpoints were altered to specifically include hip fracture or major non-vertebral fractures. Therefore, in light of the revised emphasis on hip fracture as a measure of efficacy in osteoporosis, the EMA requested a *post hoc* analysis of the effect of strontium ranelate on hip fracture incidence be provided.

---

[*] The T-scores used in TROPOS were based on local normative data available at that time. A BMD T-score of ≤-3.0 in this study is equivalent to a score of ≤-2.4 using the internationally recognised NHANES III data. Therefore, this is closely aligned to the internationally accepted definition of osteoporosis (a BMD T-score ≤-2.5 using NHANES III).

TROPOS was not designed to detect differences in hip fracture in the selected population, and consequently had insufficient statistical power to detect changes in this variable. Statistical power describes whether or not an analysis is able to robustly evaluate the effect of an intervention on a specific variable in a given study. The power is based on the population size, event frequency and effect size for the intervention. Thus, infrequent events or small populations may result in underpowered analyses, which are unreliable. It is generally accepted that a study should be powered to at least 80% to be considered as robust. TROPOS was designed to achieve a power of 90% to detect an effect size (hazard ratio) of 0.73 when evaluating peripheral fractures, which required 425 events. However, as hip fractures are a small percentage of all peripheral fractures, this trial was underpowered for the analysis of hip fractures; based on the effect size of 0.73 and the 180 hip fractures observed, the analysis of hip fractures in the whole population was powered to only 56%. Therefore, a subgroup analysis was required to provide a suitably powered investigation.

A subgroup was required in which the risk of fracture was sufficiently elevated as to enable an appropriately powered analysis to be performed. Therefore, after consultation with the EMA, the placebo arm of TROPOS was screened for the effect of three of the main risk factors for hip fracture (age, BMD and prior fracture). This approach identified a population, based solely on known risk factors, in which a robust analysis could be performed without the need for multiple exploratory evaluations.

It is generally advised that caution is exercised when evaluating *post hoc* subgroup analyses. Servier is pleased to reassure NICE that the analysis in question was the only analysis performed at that time. Therefore, Servier believes that the findings may be considered reliable and robust.

Information required

4. NICE wishes to see evidence demonstrating whether or not multiple explorations of the TROPOS data have occurred in order to identify the subgroup finally presented to EMA.

5. NICE wishes to see a full account of dealings with the EMA on the question of subgroup analysis in the TROPOS study including, but not limited to, all original documentation bearing on this question; all communication relating to the subgroup analysis; the 'day 120' questions and responses and 'day 180' meeting notes, questions and responses.

A full account of all dealings with the EMA is provided in Appendix A.

In response to a request made by the EMA in the Day 120 questions, a single subgroup analysis was performed. To select a subgroup in which the analysis would be suitably powered, the placebo arm was screened for the effects of three known risk factors on hip fracture incidence: age, BMD and prior fracture. A subgroup of women with an elevated risk of hip fracture was identified, independently of the efficacy of strontium ranelate, allowing a robust *post hoc* analysis to be performed. This subgroup was defined as women aged ≥74 years with a BMD T-score ≤-3; prior fracture was not selected as it did not significantly affect fracture risk.

Subsequently, as part of the Day 180 outstanding issues, the EMA requested specific additional supportive analyses to be performed to further explore the efficacy of strontium ranelate for the prevention of hip fractures. It is important to note that the initial subgroup analysis was carried out and reported prior to these further analyses being undertaken. The requested supportive analyses demonstrate a consistent trend towards reduced risk of hip fracture with strontium ranelate therapy. In addition, a prespecified analysis of those patients who closely adhered to the strontium ranelate regimen (the per protocol set) identified a significant reduction in the incidence of hip fracture in this population. This dose-related effect, combined with the supportive analyses described above, consistently support the findings of the primary subgroup analysis in demonstrating the efficacy of strontium ranelate for the prevention of hip fractures. Details of all supportive analyses are included in Appendix A.

Justification for the chosen subgroup

In line with contemporaneous guidelines, TROPOS was powered to identify changes in the risk of all peripheral fractures, rather than the risk of hip fracture. Overall, the percentage incidence of hip fracture was low, with only 3.1% of patients experiencing hip fracture. Therefore, a subgroup was identified in which the incidence of hip fracture was high enough to support a robust *post hoc* analysis.

This subgroup was identified by screening the placebo arm for the effect of three of the main risk factors for hip fracture incidence: age, BMD and prior fracture.

- Prior fracture was found to have no significant effect on the incidence of hip fracture in the trial population.
- Conversely, age significantly affected hip fracture rate, with women aged 74 years or over having a significantly higher risk of hip fracture than younger women. This is consistent with the real-world, epidemiological data upon which the inclusion criteria for TROPOS were based, and which identified an exponential increase in the incidence of hip fracture after the age of 74.[1] The age cut-off of ≥74 was selected to closely adhere to the inclusion criteria of TROPOS and maintain consistency with the published data on increasing risk
- Bone mineral density was also correlated with hip fracture risk: a BMD T-score ≤-3[*] (closely aligned with the internationally accepted NHANES III definition of osteoporosis) was associated with significantly higher fracture risk.

Thus, the subgroup selected for the *post hoc* analysis of hip fracture was women aged ≥74 with a BMD T-score ≤-3.

Although the selected population represents only one of several possible definitions of a higher-risk subgroup, it is based on the inclusion criteria of the study, published epidemiological data and the accepted international definition of osteoporosis. In the absence of a prespecified subgroup for analysis, this subgroup represents a higher-risk group in which a robust analysis may be performed. In addition, a number of biologically plausible explanations support the observed increase in fracture risk in this subgroup.

---

[*] A BMD T-score of ≤-3.0 in this study is equivalent to a score of ≤-2.4 using the internationally recognised NHANES III data, the reference values for the major outcome trials with bisphosphonate treatments[3,4]

These explanations may also support any interpretation of increased efficacy of strontium ranelate in this subgroup, although this is not the intention of the analysis.

Information required

7. NICE wishes to see, by reference to contemporaneous documentation at the time of the submission of the subgroup to the EMA, the justification for the age-cut off at 74.

An epidemiological study conducted in 1990 by Donaldson *et al.* investigated the incidence of fractures by age and sex, and identified an exponential increase in hip fractures in women over the age of 74.[1] This real-world finding was used as a basis for the inclusion criteria for TROPOS: this study was conducted in a defined at-risk population of women aged ≥74, or aged 70–74 with at least one additional clinical risk factor for fracture.

To ensure that a suitably powered subgroup was identified for a robust *post hoc* analysis to be conducted, a population with an elevated incidence of hip fracture was required. In order to avoid bias, the subgroup was selected by analysing the fracture incidence of patients in the placebo arm of the trial.

The cut-off values identified were both rational and representative of the real-world population. Age is a known risk factor for osteoporotic fractures and the selected cut-off, age ≥74, was a prespecified criterion for inclusion in the TROPOS trial. Furthermore, this age is consistent with real-world epidemiology findings.[1] Similarly, a BMD T-score ≤-3[*] is closely aligned with the internationally accepted definition of osteoporosis based on NHANES criteria. The validity of these cut-offs in the TROPOS population is demonstrated by the observation that both age ≥74 years and BMD T-score ≤-3 in TROPOS were associated with significantly elevated risk of hip fracture in the placebo arm of this study.

Further support for the selected age cut-off is provided by the WHO fracture risk assessment tool, FRAX. FRAX is a validated tool used to predict the risk of fracture in osteoporosis, based on real-world data. Using this tool to model the TROPOS population

---

[*] A BMD T-score of ≤-3.0 in this study is equivalent to a score of ≤-2.4 using the internationally recognised NHANES III data.

indicates that the risk of hip fracture in women aged 74 is substantially increased compared to that of women aged 70. This confirms that the subgroup of women ≥74 years is a population with an elevated risk of hip fracture, and represents a suitable cut-off point for the analysis.[§]

For reference, more details of the submission to the EMA are provided in Appendix A.

8. NICE wishes to see evidence providing a biological basis for the claim that the subgroup experiences greater benefit than the trial population overall. Please note that NICE does not regard the fact that the subgroup was accepted by the EMA as determinative.

The subgroup analysis does not necessarily indicate that strontium ranelate has a higher efficacy in the selected population, but rather that a statistically significant result was observed in an appropriately powered analysis. In line with the draft guidance available when TROPOS was designed, the study was powered to detect differences in the incidence of peripheral fractures, rather than in hip fracture specifically. Occurrences of hip fracture were recorded, but the overall incidence was relatively low. Therefore, the identification of a subgroup in which there is an increased incidence of hip fracture allows the analysis to be powered to detect a statistically significant treatment effect.

However, it is acknowledged that NICE wish to explore the idea that the subgroup results may be indicative of increased efficacy for strontium ranelate in this population, although this was not the intention of the analysis. If NICE wishes to interpret the data in this manner, then possible biological reasons for this position exist. The evidence is summarized in Appendix B. Servier accepts that NICE may wish to limit any recommendations based on this subgroup analysis to a corresponding population if this interpretation is adopted.

9. Depending on whether or not more than this high risk subgroup analysis has been performed on the TROPOS dataset (see point 4), NICE wishes Servier either to confirm

---

[§] The data on which this model is based can be supplied on request.

that the subgroup presented to the EMA was the only subgroup analysis performed on the TROPOS dataset, or to provide full details of all exploratory subgroup analyses of the TROPOS dataset and the detailed results of these analyses. Subgroup analyses should be presented demonstrating that the appropriate statistical techniques have been used to correct for multiple sub-group analyses.

The presented subgroup analysis, comprising women aged ≥74 years with a BMD T-score ≤-3, was the only subgroup analysis performed at that time. This subgroup was selected by screening the placebo arm for known clinical risk factors for hip fracture. As multiple analyses were not performed, no adjustment of significance levels is required.

Subsequent supportive analyses were performed only at the request of the EMA, and only after the initial subgroup analysis had been reported. These analyses demonstrate a consistent trend for a reduction in fracture risk with strontium ranelate therapy, thus supporting the findings of the primary subgroup analysis. All details of these analyses are included in the Appendix A.

Evidence for the biological activity of strontium ranelate is further supported by analysis of those patients who adhered closely to the recommended regimen. A greater decrease in the risk of hip fracture with treatment was observed in the per protocol set (PPS; defined as those patients with blood strontium levels above a prespecified level) than in the intent-to-treat population, demonstrating a dose-dependent effect. This sensitivity to strontium ranelate treatment implies that the reduction in hip fractures is indeed a consequence of the biological activity of strontium ranelate.

Extrapolation from the subgroup to the patient population at large

The TROPOS study robustly demonstrated that strontium ranelate is effective in the prevention of osteoporotic fractures across the broad population included in this study, as strontium ranelate gave rise to a significant decrease in the risk of all peripheral fractures compared to placebo (relative risk [RR]=0.84, p=0.04). Strontium ranelate also gave rise to a decrease in the risk of hip fracture across the whole population, although this did not reach significance (RR=0.85, p=0.058). However, this study was not powered to detect the effect of treatment on hip fracture rates. This does not reflect an inherent flaw in the study, but indicates that analysis of hip fractures across the whole population must be interpreted with caution. For this reason, the risk of hip fracture was

analysed in an appropriately powered subgroup, and a significant reduction in hip fractures was demonstrated, confirming the findings of the full analysis.

Extrapolating the findings from the subgroup analysis to a wider population is challenging. One possible approach to aid such extrapolation would be to perform further, complementary subgroup analyses in the remaining population. Indeed, the supportive analyses performed at the request of the EMA show a consistent trend in favour of strontium ranelate over placebo for the prevention of hip fractures. However, applying this approach to the population excluded from the subgroup analysis would not be appropriate, as the incidence of hip fracture is substantially reduced in this population. Based on the known population size and conservative estimates of the effect size, such an analysis would be substantially underpowered, and so would be of minimal value.

However, if the presented results are considered indicative of higher efficacy in the primary subgroup (women ≥74, BMD ≤-3), Servier accepts that NICE may wish to limit any recommendations based on this subgroup to a corresponding population, although this was not the intention of the analysis.

Importantly, the full analysis and subgroup analysis from TROPOS provide different and complementary information. Servier would encourage NICE to consider the findings from both evaluations when assessing the efficacy of strontium ranelate. The full analysis demonstrates efficacy in preventing peripheral fractures in general, whereas the subgroup analysis specifically demonstrates that strontium ranelate is efficacious in preventing hip fractures.

Information required
13. NICE wishes Servier to outline the scientific and statistical rationale for not using the efficacy data derived from the whole TROPOS clinical trial population in making recommendations that apply to this population.

In line with the guidance available at the time of inception of TROPOS, the study was powered to identify statistically significant differences in the incidence of all peripheral fractures. Hip fracture data were recorded for the full trial population and a reduction in the risk of hip fracture demonstrated (RR=0.85, p=0.058); however, the analysis was

only powered to 56%. In the case of such underpowered analyses, a lack of significance does not necessarily equate to a lack of efficacy. Rather, it indicates that the data on hip fractures across the whole TROPOS population may not be a robust representation of the true efficacy of strontium ranelate, and must be interpreted with caution. The full analysis of TROPOS robustly demonstrated a significant reduction in the risk of peripheral fracture. The subgroup analysis builds on this finding by providing a statistically powered population in which to evaluate the efficacy of strontium ranelate for the prevention of hip fractures.

14. NICE wishes Servier to present a statistical rationale for the use of a data set derived from a subgroup in a trial in preference to the use of the whole trial data set. In doing so, NICE wishes Servier to explain, from a statistical perspective, why the use of statistical analysis derived from the whole TROPOS study is not reliable or not robust.

The full analysis and subgroup analysis from the TROPOS study provide different and complementary information on the efficacy of strontium ranelate. The full trial provides robust evidence that strontium ranelate significantly reduces the risk of all peripheral fractures (RR=0.84, p=0.04) and major osteoporotic fractures (RR=0.81, p=0.031), compared to placebo, in women with osteoporosis. The subgroup analysis is powered to detect differences in hip fractures between strontium ranelate and placebo, and demonstrates a significant reduction in the risk of hip fracture (RR=0.64, p=0.046). Therefore, it is recommended that both the full and subgroup analyses are considered when evaluating the efficacy of strontium ranelate, with the full analysis providing evidence for efficacy in general, and the subgroup analysis demonstrating efficacy specifically in the prevention of hip fractures.

15. NICE would like Servier to provide statistical analyses, including central estimates of effect with confidence intervals, from the TROPOS population after the 'subgroup' dataset is removed from the overall TROPOS dataset (i.e. all women not in the subgroup) and for the population of women with a T score of-2.5 or below under the age of 74 (i.e. all women with osteoporosis under the age of 74).

Servier respectfully acknowledges the request for further exploratory analyses, beyond those requested by the EMA. However, due to the small populations and low frequency of fractures in the requested subgroups (resulting in a lack of statistical power), such analyses would be unable to provide reliable or robust estimates for the efficacy of strontium ranelate. In the case of the TROPOS population with the subgroup removed,

such an analysis would be underpowered, and therefore carry an unacceptably high chance of Type II (false negative) errors. Furthermore, a large proportion of women in this remaining subgroup may have a BMD T-score >-3, and would therefore be classed as having osteopenia, which may unduly influence such an analysis. In the case of women <74 with osteoporosis, this would be an even smaller subgroup. With such diminished patient numbers, it would be highly unlikely that an adequately powered analysis could be performed.

Conclusions

The pivotal evaluation of strontium ranelate in preventing peripheral fractures, TROPOS, robustly demonstrated that strontium ranelate is effective for the prevention of non-vertebral fractures. However, in line with the guidance available at the time of its inception, TROPOS was powered to detect significant treatment effects on the risk of peripheral fractures, and not the risk of hip fracture. Given that the guidance has subsequently been amended to specifically recommend hip fracture as a primary efficacy endpoint, the EMA requested that a *post hoc* analysis of hip fracture data be provided to support the submission. A subgroup in which a suitably powered analysis could be performed was identified by screening the placebo arm of the study, based on known risk factors for hip fracture. This analysis demonstrated a statistically significant reduction in the risk of hip fractures with strontium ranelate.

The identified subgroup was the only subgroup analysed at that time, selected using a screening process that was independent of any possible influence by the effects of strontium ranelate. Statistically and biologically justified cut-off values were chosen, that were closely aligned with the internationally accepted NHANES III definition of osteoporosis, predefined inclusion criteria and published epidemiological data. Servier understands that the extrapolation of data from subgroups to larger populations is challenging. Equally, while it is not the intention of the subgroup analysis to demonstrate higher efficacy in the selected population, the efficacy results may be interpreted in this way. In this case, it is understood that NICE may wish to limit their recommendations to a corresponding population.

It is important to note that the analyses of both the full TROPOS population and the subgroup provide different and complementary information. Therefore, it is recommended that both the full and subgroup analyses are considered when evaluating the efficacy of strontium ranelate, with the full analysis providing evidence for efficacy in general, and the subgroup analysis specifically demonstrating the efficacy of strontium ranelate in preventing hip fractures.

References

(1) Donaldson LJ, Cook A, Thomson RG. Incidence of fractures in a geographically defined population. J Epidemiol Community Health 1990;44(3):241-5.

(2) Johnell O, Kanis JA, Oden A, Johansson H, De Laet C, Delmas P, et al. Predictive value of BMD for hip and other fractures. J Bone Miner Res 2005;20(7):1185-94.

(3) Cummings S, Black D et al. Effect of alendronate on risk of fractures in women with low bone density but without vertebral fractures. JAMA 1998: 280 (24): 2077-2082

(4) McClung M, Guesens P, et al. Effect of riserdonate on the risk of hip fracture in elderly women. N Engl J Med 2001: 344 (5): 333-340