Paper 324 – Measuring prompt gamma cross-section data for Carbon target using AFRODITE clover detectors

The paper identifies an interesting and important discrepancy between experimental observations and results of Monte Carlo simulations of the same. This is of interest and should be examined.

However the title of the paper and the abstract do not correlate with each other – prompt gamma cross section data are important, and a nuclear physics research program in its own right. Similarly proton radiotherapy, and dose deposition imaging are their own research field. This paper seems confused between the two, and has very much a slant towards applications (dose imaging) rather than cross-section measurements as stated by the title.

The title of the paper has been changed to try to better reflect the focus of the paper. This paper does sit at an awkward intersection and we have tried to shift towards looking at a comparison of Monte Carlo and measured results.

In fact, in this short paper, **no cross section data** is presented. The paper as it stands presents a set of calibration data which, whilst important, is standard introductory material for a paper of this nature. The science presented here does not justify publication in its own paper.

The paper has been significantly re-written, providing a more complete description, but is limited by the maximum page count of 6. We tried to do our best within the required space and removed any reference to cross-section data.

Points of concern:

- No cross section data is presented
- This seems a calibration paper only

True, reframed paper as a comparison of measured vs simulated prompt gamma emission.

- There seems to be confusion between uncertainty, systematic error, and calibration offset.

The statements about uncertainty in the Introduction have been re-written to better reflect how uncertainty is used in proton therapy

- Compton suppression systems are discussed, but no evidence of their use in either the experimental or simulated data is given.

The Compton suppression performance has been included in the simulation study. A statement was added in Section 6.2

- The GEANT model of the APHRODITE array has incomplete discussion. It is not clear if the whole array is modelled, with mechanical structure and vault, the whole array, or just one clover. The BGO Compton scatter suppression is omitted from this discussion.

The whole AFRODITE array was modelled and used for all simulation study. A statement was added to Section 3.

- Figures 4 and 5 are not discussed in the text. Further detail is required.

Figure 3 and 4 were removed and Figure 5 was replaced with a Co-60 spectra.

- "The gap between the experimental and simulated spectra is due to the lower efficiency of the simulated gamma sources." Makes no sense, and requires a much fuller discussion. The "gap" refers to a systematic offset, which may or may not be linear (difficult to tell on a log plot). "The lower efficiency" could mean anything….

The gap between the simulated and measured calibration spectra no longer exists. A better normalization was used to compare the two spectra, results in a much smaller difference (about 15%) between the two.

- It is not clear to this reviewer what "Geant 4 simulation **normalised**" means. One would have thought that normalisation would remove systematic offsets between data sets, but this is clearly not the case.

In the new figures 3, 4, 5, the normalization is explained in the caption for each figure.

- It is interesting to note that simulation underpredicts the experimental values (fig 5), but has some gamma lines which are not in the experimental data. It would be interesting to investigate these. It is normally the case that simulation *overpredicts* efficiency and that experimental data has more gamma lines than expected.

Yes, this is true, in the updated figure 4, the simulation still underpredicts the experiment, but this can be attributed to the simulated Compton-suppression, which is working a little too good.

In the updated Figure 5, the simulation does overestimate the gamma spectra compared to the measured spectra.

- It is not clear what figure 8 is presenting. The fitted line is misleading, or badly described.

Figure 8 has been removed and the absolute detector efficiency is referred to in Section 6.3

- Section 4.3 changes the simulation from describing the experimental setup (which is not explained in enough detail), to a new scenario. This is worrying.

The only difference between the experimental and simulated setups is the target thickness. The target thickness was increased for the simulation in order to get good results in a decent amount of time (days instead of weeks). The difference in thickness was corrected for in the final simulated result (new figure 5) and discussed in Section 6.3

- Figure 9 presents a comparison between experimental and simulated data in this case. Systematic offsets are observed and not explained. In this case simulation over predicts efficiency in contrast to the above data, and introduces features which are not observed experimentally. Some features on both plots have offsets in count rate, and in energy.

Figure 9 has been replaced with the updated figure 5. An updated normalization has been used and there is better alignment between the two spectra making it easier to distinguish differences. The simulation still exhibits small difference (and an overall over-estimation) from the measured spectrum, but we are primarily concerned with the 4.438 MeV peak.

- Fig 9 circle highlighting 4.438 MeV is not discussed.

Specific details regarding the 4.438 MeV peak has been added to Section 6.3

- The conclusion is loose and open to interpretation. How does one quantify a "favourable comparison"?

Wording in the conclusion has been re-written.