Discussion on best practice in labelling MC predictions in EXP plots

Fabio & Mike

Prologomena

- Our work in recent years has lead to an increase in:

  - Number of MC tools (often in combination) available to the EXP community.

  - Number of methods and techniques.

  - Accuracy and precision of predictions.

Prologomena

- This is obviously great, but it has also lead to a number of unexpected/unwanted side effects, among which:

  - Complexity of the MC offer which needs to be explained to users, including the fact that a given MC tool can provide predictions at different accuracy.

  - The definition of accuracy itself has been enriched from a simple LO, NLO and so on, to LO,NLO,NLO+PS, NNLO+PS, MC@NLO, S-MC@NLO, POWHEG, MiNLO, MEPS@LO, MEPS@NLO, ULOPS, UNLOPS, MLM-Kt, FxFx, NLO+PS (QCD+EW),…..

 - Inability of the EXP community to correctly handle this…

Example #1

arXiv:1505.06520
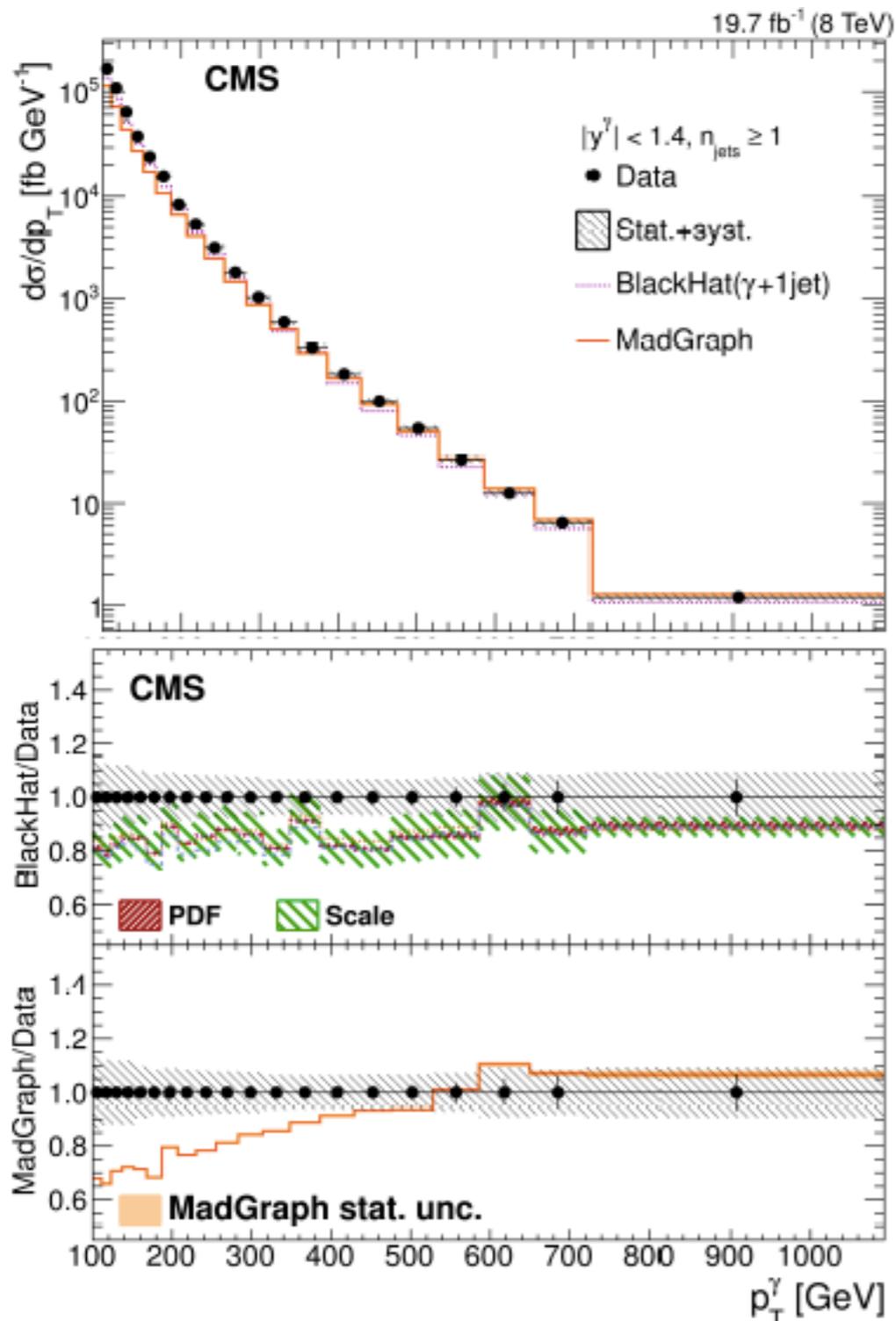
Guess what are the
theory curves here!

Example #1
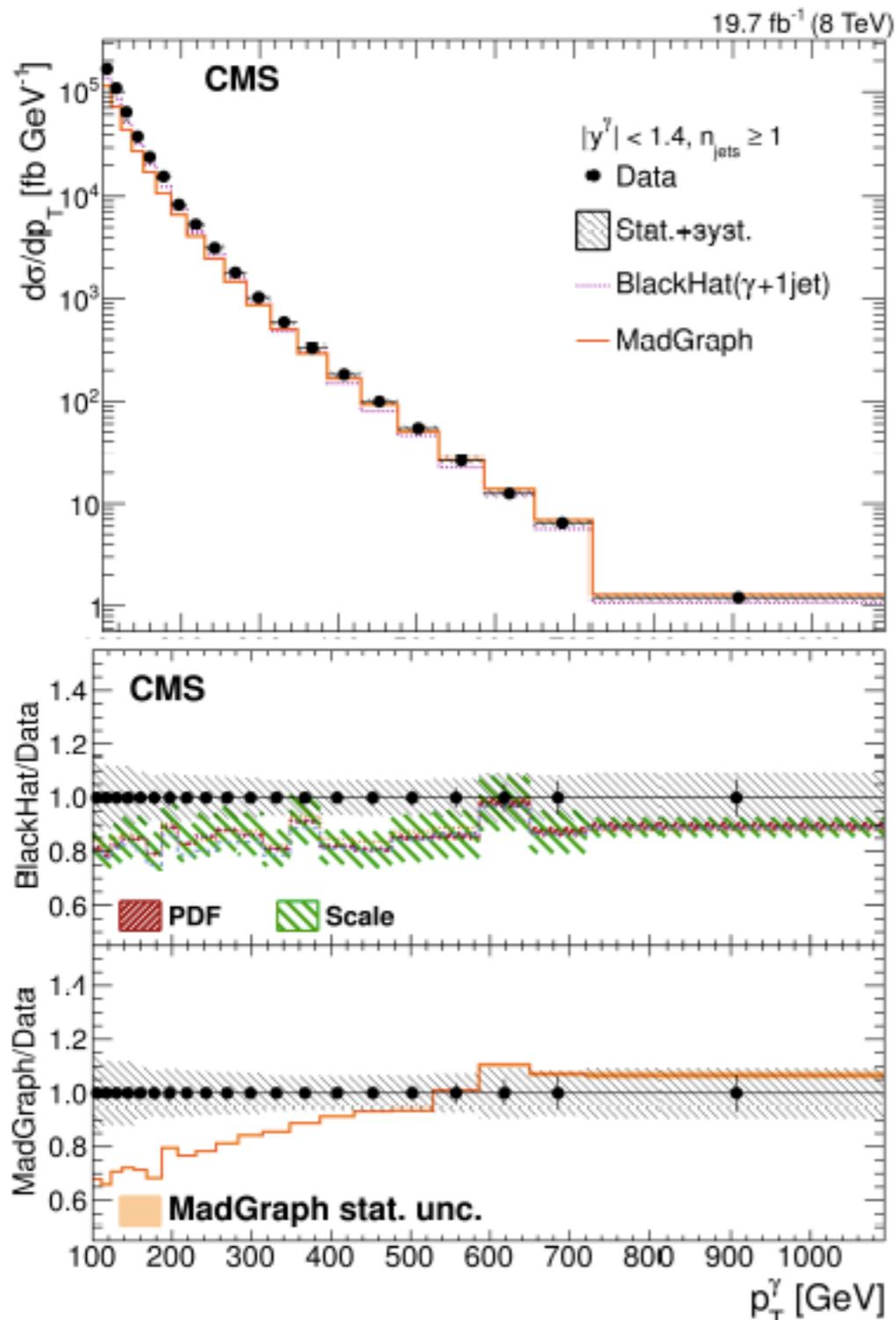
arXiv:1505.06520

Answer:

"BlackHat"= NLO parton level fixed order actually obtained in the SHERPA framework.

"MadGraph"= MLM-KT merged LO multiparton predictions (up to 4 partons) in the MG5_aMC framework + Pythia6.

Comments:

Here the name of the code was used also to imply a given accuracy, actually creating confusion.

Example #2

arXiv:1505.06520
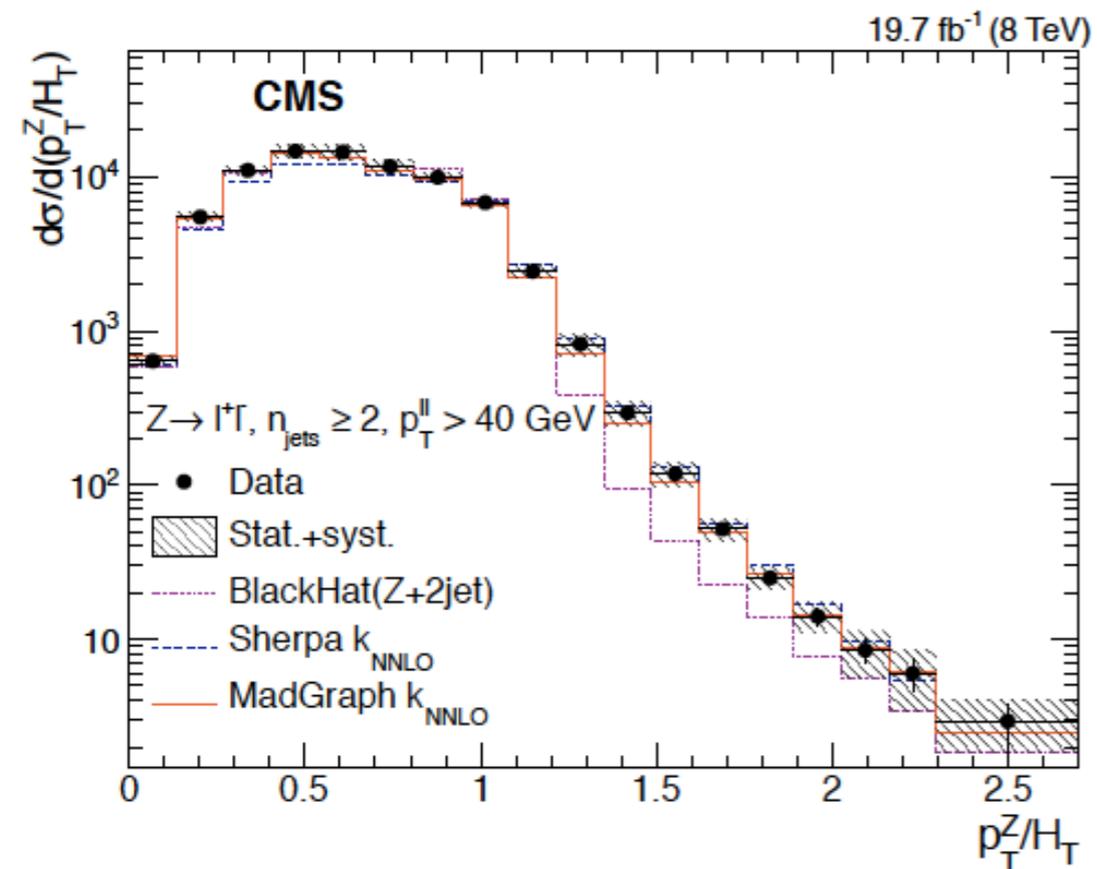
Here:

"BlackHat"= NLO parton level fixed order actually obtained in the SHERPA framework.

"MadGraph"= MLM-KT merged LO multiparton predictions (up to 4 partons) in the MG5_aMC framework + Pythia6. Normalised to NNLO.

"Sherpa"= CKKW merged LO multiparton predictions (ME+PS). Normalised to NNLO.

# Example #3

CMS-PAS-SMP-14-021

pp →γγ
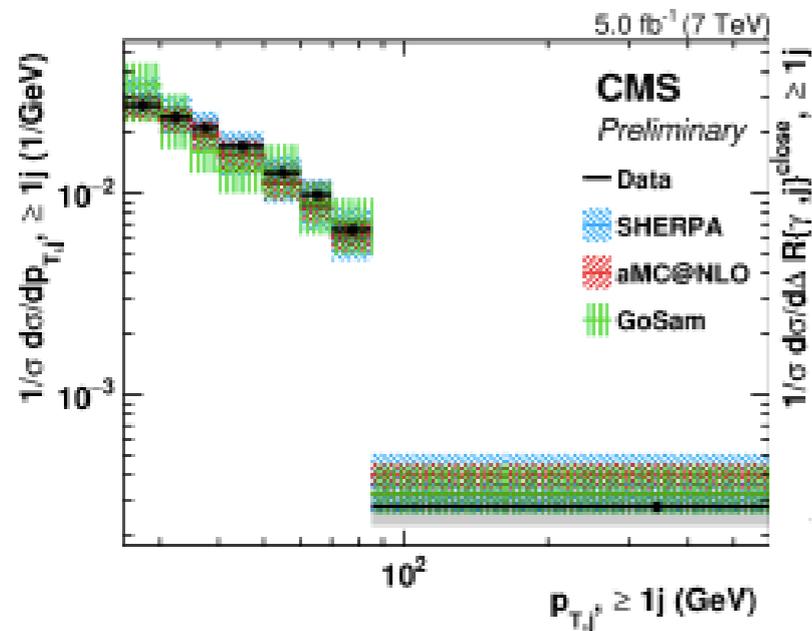
Guess the what are the labels now….

# Example #3

CMS-PAS-SMP-14-021

pp →γγ
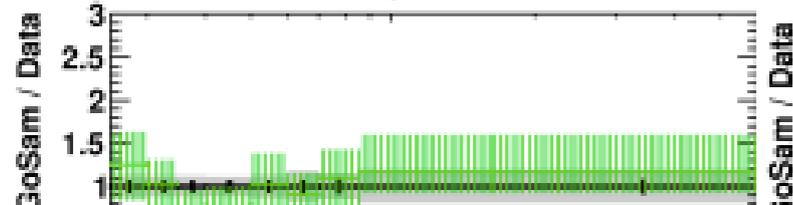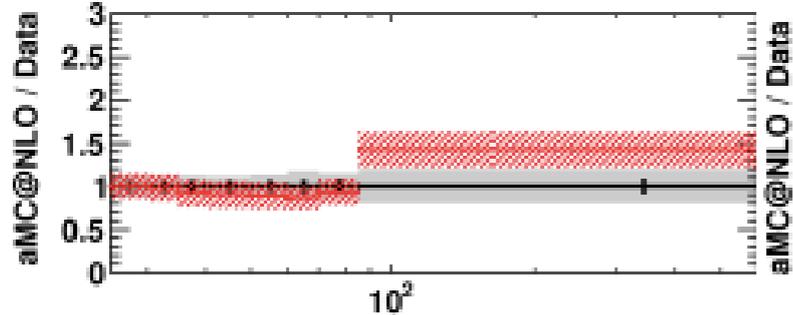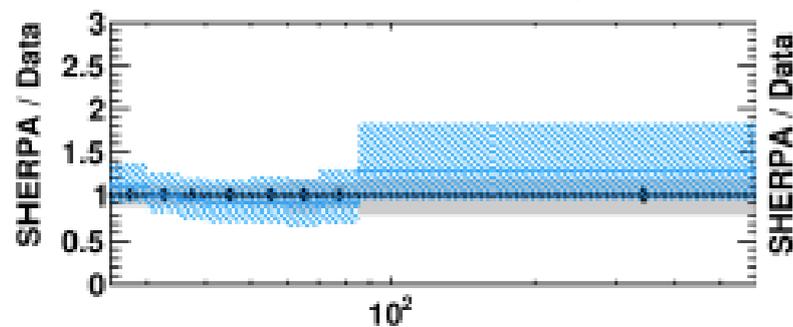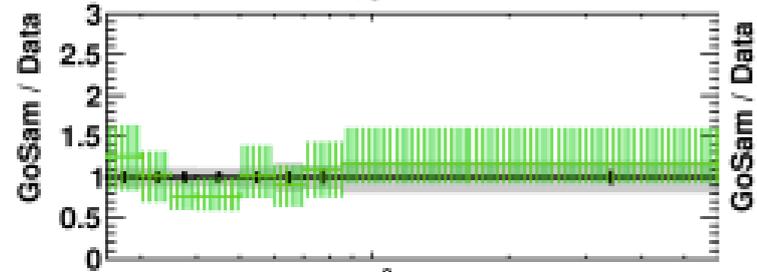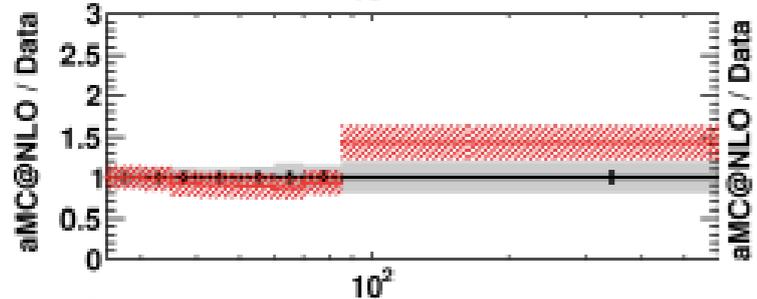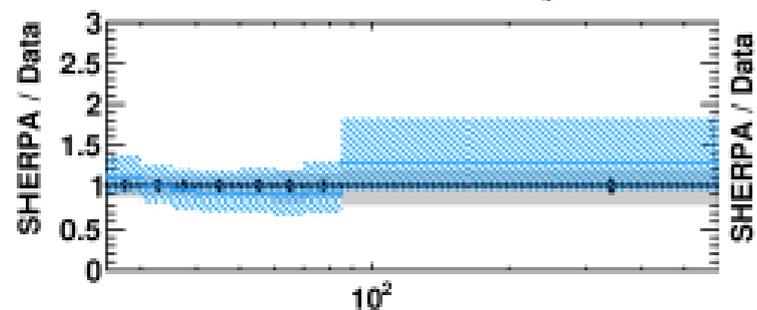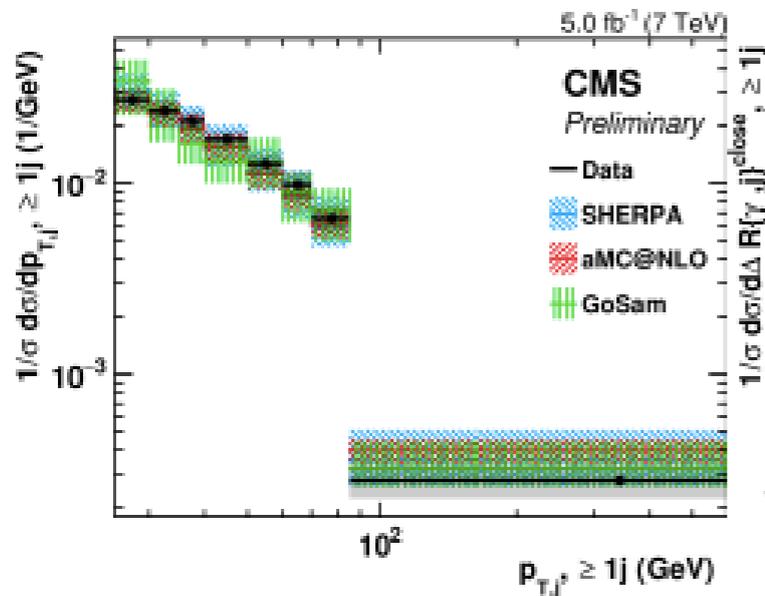
Here:

"Sherpa"= LO ME+PS (up to 3 partons+gg fusion) (I think).

"aMC@NLO"= NLO+Pythia6

"GoSam" = NLO fixed order, parton level, with GoSam loops in the MadGraph/MadDipole framework. Note that GoSam is a loop provider in general not a MC…

Comments

- Obviously, all the relevant information to reproduce a sample cannot be contained in the caption of a plot. The full information needs to be included in the text (which is not very often the case).

- Nowadays giving the names of codes such as MadGraph (or MG5aMC), Sherpa and Herwig does not say anything about the accuracy of a prediction. This kind of labelling should be deprecated as it leads to statements like Code 1 does better than Code 2, which in general does not make any sense (if predictions are at different order for a given observable)

MCnet Guidelines

MCNET GUIDELINES

for Event Generator Authors and Users

PREAMBLE

This generator has been developed as part of an academic research project and is the result of many years of work by the authors. Proper academic recognition is a requirement for its continued development.

MCnet Guidelines

2) The program and its physics should be properly cited when used for academic publications

2.3) When several programs are combined, they should all be mentioned, commensurate with their importance for the physics study at hand.

2.4) To make published results reproducible, the exact versions of the codes that were used and any relevant program and parameter modifications should be spelled out.

Open questions

- Can we (MCnet) do something to improve this situation?

- Can we agree to always label the MC curves with the accuracy first and then the names of the codes?

accuracy of the simulation+ name_of_the_code

Possible caveat: the accuracy of a prediction depends on the specific observable plotted. While this is certainly true, I believe that the level of possible misunderstanding/misuse would be lower than using the name of the code…

- If the accuracy is the same, it can be factored out in the global label of the picture.

- Can we agree on some kind of naming scheme or guidelines?

Proposals?

- Use a few kind of standardised labels:

  - LO, NLO, NNLO (parton level predictions at a given accuracy)

  - LOPS, NLOPS (or LO+PS, NLO+PS) (particle level predictions matched). In the case of external PS program is employed, the name of the program should be used instead of the generic PS.

  - MEPS@LO, MEPS@NLO, MLM-KT, CKKW , FxFX : can we choose a neutral name and specify the details of the method in the main text?

- When emphasis of the comparison is on specific issues those should be clearly specified (shower, underlying events, codes, and so on). If the comparison is consistent (for example different predictions at NLO+PS), then this could be specified as a global label and the names of the codes can be used.

- When NLO can be both QCD and EW, the order, such as NLO (QCD+EW) should be clearly specified.

- In case no space is available in the plot (I have heard this as a motivation several times) and computations are not at the same accuracy, preference should be given to the accuracy than the name of the codes.