Peer Review File

Article Information: http://dx.doi.org/10.21037/gs-20-548.

Reviewer A

The article is very interesting and well written. The methodology is correct and the study is well organized. The numerical sample is good and the conclusions are fair. I think the article can be published in the current version

Reply: Thanks for the reviewer's kind comments.

Reviewer B

1. **Material and Methods. Laboratory procedures to measure serum calcium are lacking.**

Reply 1: Thanks for the reviewer's useful comment.

Changes in the text: we added "Total serum calcium was measured on the AU5800 analyzer (Beckman Coulter, USA) (normal range 2.20–2.65 mmol/L)" (see Page 7, line 110-112).

2. **Corrected calcium is a preferred measurement to assess hypocalcemia in the evaluation of patients after thyroidectomy. Authors should state why they have not used this analytical parameter.**

Reply 2: We agree with the reviewer's professional suggestion. However, corrected calcium needs the measurements of serum total calcium and albumin according to the equation (Corrected total calcium = measured total calcium + 0.8 (4.0 − serum albumin)). However, serum albumin was not routinely measured during follow-up. Thus, we did not use corrected calcium to assess hypocalcemia.

Changes in the text: we added "Albumin corrected total calcium was not investigated because serum albumin was not routinely measured during follow-up" in Methods (see Page 7, line 112-113).

3. **Pag. 6, lines 95-98. It is difficult to understand that a PTH assay could measure with precision serum concentrations as low as 0-1 pg/ml. The authors should clearly state the limit of detection (LOD), limit of blank (LOB), limit of**

quantification (LOQ), analytical sensitivity, functional sensitivity of the PTH assay they have used. The methodology used to establish the functional sensitivity of their assay at their own laboratory has also to be clearly commented in the section Material and methods.

Reply 3: We thank the reviewer's kind suggestions. The reviewer's comments are very professional. We read the instructions for the kit for determination of parathyroid hormone (chemiluminescence method) and consulted our colleagues in the laboratory. The limit of detection and limit of blank of PTH assay are both 1 pg/mL (0.1 pmol/L). The limit of quantification and functional sensitivity are both < 4 pg/mL. The measurement of PTH concentration > 1 pg/mL was achieved by so-called Regular Mode on the DXI 800 Immunoassay System. While the measurement of PTH concentration < 1 pg/mL was achieved by so-called Intraoperative Mode on the DXI 800 Immunoassay System. In the Intraoperative Mode, an equation derived from the relationship between the known standard analyte concentrations and their relative light units. The relative light unit and the corresponding PTH value was calculated through the equation when PTH concentration was < 1 pg/mL. Honestly, the value < 1 pg/mL was not accurate and repeatable based on my colleagues' experience in the laboratory. That is to say, 0, 0.1, or 0.2 pg/mL might not be much different from 0.3, 0.4 pg/mL. However, thyroid and parathyroid surgeons in our hospital wanted to know the specific PTH value other than simple "< 1 pg/mL" in the Regular Mode and understood the imprecision of the assay by the Intraoperative Mode when performing parathyroid surgery in the past ten years. In our experience, even an extremely low serum PTH concentration may, to some degree, indicate a better parathyroid function than zero. The reviewer's suggestion is right. To avoid unnecessary misunderstanding and make readers better understand our results, we followed the reviewer's suggestion and added information about the imprecision of the assay of serum PTH concentration was < 1 pg/mL in the Intraoperative Mode on the DXI 800 Immunoassay System.

Changes in the text: we added above-mentioned information in Methods (see Page 7, line 101-110).

4. **Furthermore, the measurements of the imprecision of the assay (coefficients of variation) at different levels of serum PTH concentration should be also stated in the paper, especially at the low and very low levels of PTH**.

Reply 4: Thanks for the reviewer's suggestion. This comment is also about the imprecision of the assay of PTH concentration < 1 pg/mL. It is true that the imprecision of the assay at different levels of serum PTH concentration should be confirmed. In our laboratory, total imprecision was evaluated in the Regular Mode if the serum PTH concentration is not low (see Table 1). If the serum PTH concentration is below 1 pg/mL, the imprecision of the assay was evaluated the Intraoperative Mode (see Table 2).

Honestly, the authors of this article are not much familiar with the meanings of these laboratory terms even after consulting our laboratory colleagues. We have learned much from the reviewer. We agree with the reviewer's suggestion that the imprecision of the assay of PTH concentration <1 pg/mL in the Intraoperative Mode should be mentioned in the manuscript for those potential readers, who may also be not familiar with those laboratory terms. Evaluation of the imprecision of the assay at different levels of serum PTH concentrations with tables will make the manuscript much complicated. And the potential readers may focus more on the truth of inaccurate results than the evaluation process. So the response to this comment is similar to Comment 3.

Table 1. Imprecision - Regular Mode

Human EDTA	Pg/mL	Within group	Between groups	Total imprecision
Serum sample	Pmol/L	%CV	%CV	%CV
Level 1	12.1 (1.3)	2.6	5.8	6.4
Level 2	144 (15.3)	1.6	3.2	3.6
Level 3	1439 (152.5)	2.2	2.8	3.5

Table 1. Imprecision - Intraoperative Mode

Human EDTA	Pg/mL	Within group	Between groups	Total imprecision
Serum sample	Pmol/L	%CV	%CV	%CV
Level 1	11.4 (1.2)	6.8	8.1	10.6
Level 2	144 (15.3)	2.8	3.3	4.4
Level 3	1433 (151.9)	3.2	3.0	4.4

Changes in the text: we added above-mentioned information in Methods (see Page 7, line 101-110).

5. **The authors should clearly state that with the immunoassay system from Beckman it is possible to differentiate serum PTH concentrations between 0 and 1 pg/ml with appropriate precision. Interestingly, according to the manufacturer's insert, the Beckman Coulter assay has an LOD of 3.2 pg/ml and an LOQ of 4.89 pg/ml. In other words, how could the authors be sure that serum PTH concentrations of 0.2 or 0.4 pg/ml are different from 0. Have they considered the imprecision of their assay?**

Reply 5: We thank the reviewer's valuable comments. Obviously, the reviewer spent much his valuable time reading our manuscript and the instructions. We really appreciate the reviewer for his efforts. Similarly, this suggestion is also about the imprecision of the assay of serum PTH concentration < 1 pg/mL. The reviewer is correct. As mentioned above, the DXI 800 Immunoassay System is able to measure serum PTH concentration < 1 pg/mL with a specific value in the Intraoperative Mode. Beckman's test kit and its associated calibrators are calibrated regularly. Calibration curves are determined by six known standard concentrations (S0-S5 respectively). The calibration process is to test the samples with known standard analyte concentrations just like testing patients' samples. The relative light unit (RLU) is obtained by testing the calibrator. The mathematical relationship between the measured RLU and the known analyte concentration can be used to establish the equation. The quantitative analyte concentration of PTH can be obtained by introducing the RLU of patient sample measusred into the equation. In the Intraoperative mode, the RLU of some patient is very low, and a numerical value can also be obtained when it is brought into the equation. For example, if the data is <1 pg /mL, it may be inaccurate. Clinicians will be informed of the imprecision of the assay of the Intraoperative Mode. In the Regular Mode, it will be read as <1 pg /mL without a specific value.

Truly, we are not sure that serum PTH concentrations of 0.1, 0.2 or 0.5 pg/ml are different from 0. Though in our experience, an extremely low serum PTH concentration with a value may, to some degree, indicate a better parathyroid function than zero, the imprecision of PTH values < 1 pg/mL is the truth according the instructions of the DXI 800 Immunoassay System. So the response to this comment is also similar to Comment 3 and 4. Readers will be informed of the imprecision of PTH

measurements and judge the results and conclusions in this manuscript by themselves.
Changes in the text: we added above-mentioned information in Methods (see Page 7, line 101-110).

6. It is important for clinicians to understand that it is possible to detect PTH concentrations as low as 0.1 or 0.2 pg/ml and that this has clinical relevance. Authors should specify how they have achieved this in their laboratory.

Reply 6: Thanks for the reviewer's kind comments. The measurement of PTH concentration <1 pg/mL was achieved through the equation derived from the relationship between the relative light unit and standard analyte concentration by the Intraoperative Mode on the DXI 800 Immunoassay System. Beckman's test kit and its associated calibrators are calibrated regularly. As above-mentioned calibration process of the DXI 800 Immunoassay System, calibration curves are determined by known standard concentrations. The relative light unit (RLU) is obtained by testing the calibrator. The mathematical relationship between the measured RLU and the known analyte concentration establishes the equation. The quantitative analyte concentration of PTH can be obtained by introducing the RLU of patient sample measured into the equation. In the Intraoperative mode, a numerical value can also be obtained when the RLU of some patient, which is very low, is brought into the equation. However, the clinical relevance of these measurements should be carefully interpreted by doctors. It will make the manuscript much lengthy if we try to clarify how very low serum PTH concentrations were achieved in our laboratory. Therefore, we just added some sentences to inform the reader clearly that the measurement method was imprecise when PTH concentration is very low.

Changes in the text: we added "However, these measurement results of serum PTH concentrations <1 pg/mL in the Intraoperative Mode were imprecise and not repeatable. Thyroid surgeons in our hospital would carefully interpret these results based on actual clinical scenarios." in Methods (see Page 7, line 107-110).

7. Pag. 7, lines 105-111. Protracted hypoparathyroidism was defined as PTH <12 pg/ml at 2 months after thyroidectomy. However, there are patients with serum PTH higher than 12 pg/ml who need calcium and vitamin D supplements at this time. How did they classify these patients? Some of them may reach 12 months of

follow-up in the same situation and they are classified as permanent hypoparathyroidism according to the classification stated by the authors.

Reply 7: The reviewer raises an interesting concern. We agreed that patients with serum PTH >12 pg/mL who still needed calcium and vitamin D supplements 2 months after surgery could be diagnosed protracted hypoparathyroidism. It is similar with the definition of permanent hypoparathyroidism. However, here we introduced the concept of protracted hypoparathyroidism in the current study in order to investigate the speed of recovery of damaged parathyroid function at the 2-month time point. In addition, we routinely prescribed calcium supplements to patients if his PTH is below 12 pg/mL on post-operative day 1. In other papers, some researchers chose 2-week, 4-week or 6-week time points to examine the extent of parathyroid function damage and the tendency to develop into permanent hypoparathyroidism based on the serum PTH values (references: Lorente-Poch L, Sancho JJ, Munoz-Nova JL, et al. Defining the syndromes of parathyroid failure after total thyroidectomy. Gland Surg 2015;4:82-90.; Sitges-Serra A, Ruiz S, Girvent M, et al. Outcome of protracted hypoparathyroidism after total thyroidectomy. Br J Surg 2010;97:1687-95). Whether patients with protracted hypoparathyroidism needed calcium supplements or had a greater chance to develop into permanent hypoparathyroidism was not the focus of the current manuscript. Transient hypoparathyroidism is usually acceptable for patients and doctors in the clinical scenarios, thus definition of protracted hypoparathyroidism here is not very important and so strict. Thus, we define protracted hypoparathyroidism simply according to the serum PTH value, regardless of hypocalcemia-related symptoms or the need of calcium supplements. The reviewer's comment is useful. We added some information to clearly define protracted hypoparathyroidism here.

Changes in the text: we added "regardless of hypocalcemia-related symptoms or the need of calcium supplements." (see Page 8, line 123-124), and deleted "which seemed to also have a high chance to develop into permanent hypoparathyroidism" (see Page 12, line 217-218).

8. **Results. Section kinetics of SDPF-D1 during the following 12 months after thyroidectomy. Pages 8-9, lines 135-144. This text should be notably shortened. It is not necessary to repeat the information in text and figure 2**.

Reply 8: Thanks for the reviewer's useful suggestion. We shortened unnecessary

content.

Changes in the text: we deleted some sentences (see Page 9, line 155-157 and 158-161).

9. Page 9, lines 145-151. Same consideration as before. Avoid to repeat information.

Reply 9: Thanks for the reviewer's kind comment. We have deleted duplicate and unimportant information following the reviewer's suggestion.

Changes in the text: we deleted some sentences (see Page 10, line 167-171).

10. Table 1. It is not adequate to consider protracted hypoparathyroidism as a risk factor for permanent hypoparathyroidism. It is expected that all patients with permanent hypoparathyroidism have had previously hypoparathyroidism at discharge of surgery and protracted hypoparathyroidism at 2 months.

Reply 10: Thanks for the reviewer's intelligent suggestion. We deleted the item of "protracted hypoparathyroidism" in Table 1, Table 2 and Table 3.

Changes in the text: we deleted the item of "protracted hypoparathyroidism" in Table 1, Table 2 and Table 3 (see Table 1, Table 2 and Table 3).

11. Table 1. The precision to the thousandth in the concentration of calcium expressed in mmol/l is surprising. They should explain how they achieve this or omit this precision.

Reply 11: Thanks for the reviewer's useful comment. Raw data from our laboratory was accurate to the percentile. However, the data generated by statistical analysis software was accurate to the thousandth. We did not notice this question. Thank the reviewer for pointing it out. We chose to omit this precision.

Changes in the text: we modified some data, and the precision to the thousandth was changed to the percentile (see Table 1, Table 2 and Table 3).

12. Table 2. There some percentages with 3 decimal places. It is not necessary.

Reply 12: Thanks for the reviewer's comment. We modified the data following the reviewer's suggestion.

Changes in the text: we modified percentages with 3 decimal places to percentages

with 1 decimal place (see Table 2 and Table 3).

13. Table 2. The variable postoperative PTH on day 1 is different between the two studied groups. This is by definitions of the groups.

Reply 13: Thanks for the reviewer's comment. To avoid misunderstanding, we deleted the P value of this item.

Changes in the text: we deleted the P value of this item (see Table 2).

14. Table 3. It is unnecessary to put 3 decimal places in the percentages.

Reply 14: Thanks for the reviewer's comment. We modified the data.

Changes in the text: we modified percentages with 3 decimal places to percentages with 1 decimal place (see Table 3).

15. According to the data reported by the authors the prevalence of permanent hypoparathyroidism is 1.2% (11 out of 949 patients with total thyroidectomy). This is a surprisingly low value in comparison with other surgical series. Recent reports on multicenter studies (Díez et al. Endocrine 2019;66:405) and large cohort studies (Annebäck et al. Ann Surg 2020) have shown prevalence values of permanent hypoparathyroidism very higher than that reported in this manuscript. This deserves some comment.

Reply 15: The reviewer is much professional, and really raises an interesting concern. Permanent hypoparathyroidism rate varies among studies. It is a big issue to figure out the underlying reasons. There are many factors resulting in different permanent hypoparathyroidism rates reported such as lack of clear definitions of permanent hypoparathyroidism, different follow-up duration (6 months or 12 months), different indications for surgery and case mix, incomplete follow-up, different sample size, difference in surgeons' experience or skills (see references: ***1.*** Orloff LA, Wiseman SM, Bernet VJ, et al. American Thyroid Association Statement on Postoperative Hypoparathyroidism: Diagnosis, Prevention, and Management in Adults. Thyroid 2018;28:830-41. ***2.*** Lorente-Poch L, Sancho JJ, Munoz-Nova JL, et al. Defining the syndromes of parathyroid failure after total thyroidectomy. Gland Surg 2015;4:82-90. ***3.*** Mehanna HM, Jain A, Randeva H, et al. Postoperative hypocalcemia--the difference a definition makes. Head Neck 2010;32:279-83.). Our permanent hypoparathyroidism rate (1.2%) is similar with other researchers' results (permanent hypoparathyroidism rate: **1%,** see Zhang L, et al. J Clin Endocrinol Metab 2012;97:1250-7.;

permanent hypoparathyroidism rate:**1.9%,** see Ritter K, et al. J Surg Res 2015;197:348-53.; permanent hypoparathyroidism rate: **0-3%,** see Edafe O, et al. Br J Surg 2014;101:307-20.). Moreover, we downloaded and reviewed the two references the reviewer provided. The permanent hypoparathyroidism rate of 14.5% reported by Díez et al and 12.5% reported by Annebäck et al at 12 months are very high. We don't know why permanent hypoparathyroidism rates in many reports are higher than that reported in this manuscript. Was it due to more parathyroid tissue found at pathological specimen in the study of Díez et al? We suppose the high rates of these two studies may mostly be related to surgeons with different experience and surgical volume ("the expertise of the surgical team is related to recovery of parathyroid function" reported by Díez et al in their article). In our experience, it is extremely important to identify and preserve parathyroid glands in situ for preventing permanent hypoparathyroidism. For example, increased PTGs autotransplanted and removed, and significantly low PGRIS score were detected in SDPF-D1 in our study. In the two papers the reviewer mentioned, thyroidectomies were performed by different surgical teams across many years. As time goes on, thyroid surgeons can become more and more experienced and have more understanding of anatomy of thyroid and parathyroid glands. And this helps to decrease permanent hypoparathyroidism rate.

In the current study, all operations were performed by an experienced surgeon (H.L.) with a volume of more than 800 thyroidectomies per year. It may partly account for the lower permanent hypoparathyroidism rate reported in our manuscript. Report bias was also one of the reasons. Many patients, who might have permanent hypoparathyroidism, did not adhere to follow-up and were not included in the current study. Comparison of different permanent hypoparathyroidism rates was not the focus of the current study, thus we did not elaborate on this issue in the Discussion.

Changes in the text: we added "All operations were performed by an experienced surgeon (H.L.)." in Methods (see Page 6, line 90-91).

Moreover, "The experienced surgical team is essential to accurately identify and preserve PTGs in situ. The permanent hypoparathyroidism rate reported in this study was much lower than those reported in many reports. In the current study, all operations were performed by a very experienced thyroid surgeon (H.L.) with a high volume per year. This might partly account for the low permanent hypoparathyroidism rate reported in our study. Moreover, report bias was also one of

the reasons. Many patients from distant areas, who might have permanent hypoparathyroidism, did not adhere to follow-up and were not included in the current study" was added in Discussion (see Page 14, line 258-266). Those two references (Díez et al. and Annebäck et al.) were also added to Reference section. (see Page 20, line 377-382).

16. Discussion. Page 12, lines 209-211. The authors did not comment on the noteworthy sensitivity and precision of their assay.

Reply 16: Thanks for the reviewer's comment. We should emphasize the limitation of sensitivity and precision of our assay of serum PTH concentration < 1 pg/mL.

Changes in the text: we added "Furthermore, the poor sensitivity and imprecision of the assay of serum PTH concentration < 1 pg/mL in the current study still needs further discussion" in Discussion (see Page 13, line 228-230).

Reviewer C

The authors propose an interesting retrospective review of patients with POD1 low PTH and the timing underlying their recovery.

The article is complete and the database appears solid, although some revision would be advisable in my opinion in order to improve the overall quality of the work and its internal and external validity.

1. First and foremost the title is a little bit misleading, the authors seem to refer to the timing of PT function and not about the kinetics which is an inappropriate use of terminology

Reply 1: Thanks for the reviewer's kind comment. We modified the title following the reviewer's suggestion. We changed "kinetics" to "recovery" in the title.

Changes in the text: We changed "kinetics" to "recovery" (see Page 1, line 3; see Page 5, line 76, 78; Page 9, line 150). Furthermore, "Kinetics" was changed to "time-related changes" (see Page 3, line 34; Page 15, line 269).

2. Through the entire text the expression "some 122" with different numbers. This expression induces a general sense of inaccuracy which makes the text weaker. Simply state the number (with its denominator whenever you feel it is need, as you did in the results section of the abstract).

Reply 2: Thanks for the reviewer's useful comment. We modified the expression of "some …" with different numbers following the reviewer's suggestion.

Changes in the text: we deleted "some" and state the exact number (see Page 3, line 43; Page 8, line 144,148; Page 21, line 388).

3. The results section (both abstract and main text) is a little overcomplicated and could benefit from further restructuring.

Reply 3: Thanks for the reviewer's suggestion. We deleted some unimportant and repetitive content in the results section (both abstract and main text).

Changes in the text: we deleted some sentences in Abstract (see Page 3, line 46-47, 49-51) and in Results (see Page 9, line 155-161; Page 10, line 167-171).

4. Most of my doubts come from the handling of data. The authors rightfully state that they differentiated normally and non-normally distributed data but they provide no information on how they checked normality (kolmogorov-smirnov would do)... consequently data should be presented coherently in the text and in tables, with normally distributed data presented with mean and SD and non-normal distributions with median, IQR and range.

Reply 4: Thanks for the reviewer's suggestion. Here, D'Agostino & Pearson omnibus normality test was used to check normality in GraphPad Prism 5 in our study. We added this important information in the section of Statistical analysis.

Changes in the text: we added "D'Agostino & Pearson omnibus normality test" in Methods (see Page 8, line 135-136). "Normal distribution of the data was determined using the De Agostino–Pearson omnibus normality test" was added in the explanatory legends of Table 1, Table 2 and Table 3 (see Table 1, Table 2 and Table 3).