Soluble thyroid tumor markers – old and new challenges and potential solutions

Stefan KG Grebe

Introduction

The New Zealand Institute of Medical Laboratory Science (NZIMLS) kindly invited me in 2012 to their annual scientific meeting in Wellington. When I accepted that invitation I agreed to subsequently provide a review article, based on the content of one of my talks, which was entitled "Thyroid Tumor Markers – new challenges and solutions". Regrettably, it has taken me much longer than anticipated to honor this obligation, but I believe that this much belated review should still be relevant to clinical laboratory scientists, clinical and biochemical pathologists, oncologists and endocrinologists.

For most of the 20th century, thyroid cancer was considered a relatively rare and typically non-lethal cancer. Interest in thyroid cancer related laboratory testing was consequently relatively low. However, during the last 20-25 years, incidence rates of thyroid cancer have risen exponentially in nearly all countries, while the already low cause-specific mortality rates have further declined. Consequently, it can be predicted that by the end of the 21st century thyroid cancer will be the 3rd or 4th most common cancer diagnosis in living cancer patients in most parts of the world. Since 15-40% of patients will suffer (mostly loco-regional) recurrences through their (usually) long life with their disease, there is an exponentially growing need for cost effective and accurate laboratory tests that detect (or predict) recurrence, in order to minimize treatment-related morbidity and costs. The current mainstay tumor marker used for this purpose is thyroglobulin (Tg). However, Tg is an imperfect tumor marker. Moreover, its utility might be further eroded in the future by nascent trends toward less extensive initial tumor surgery.

In this review, we will examine potential solution to the known shortcomings of Tg and will also explore ways to address some of the anticipated future issues related to less radical initial treatment.

We will start with a brief overview of the types of thyroid cancer that will be covered in this paper. Next, we will expand on the reasons why we the need thyroid tumor markers, followed by a summary of the strengths and limitations of the main current marker, Tg. Finally, we will discuss potential solutions for the current (and the anticipated future) limitations of Tg testing.

The types of thyroid cancers that will be discussed

The thyroid gland consist of three major components: follicular cells (thyrocytes), which produce thyroid hormones, parafollicular cells (C-cells), which produce calcitonin, and stromal tissue, which consists chiefly of (quite abundant) blood vessels, lymphatic vessels and connective tissue. While any of these structures - or metastases from malignancies of other organs - can give rise to malignant thyroid neoplasms, well over 90% arise from the follicular epithelium, and it is these follicular cell-derived carcinomas, which have been increasing in incidence every year during the last 2-3 decades (1-3). Consequently, this review will focus on this group of neoplasms. In addition, we will limit our coverage to those follicular cell-derived thyroid cancers that retain at least a modicum of differentiation, because they outnumber the vastly more aggressive undifferentiated (anaplastic) thyroid carcinomas (ATC) by a factor of 20 or more. In addition, median survival of patients with ATC is measured in months, rather than years, obviating any significant need for long-term tumor marker use during follow-up.

The case for tumor marker use in the follow-up of differentiated thyroid cancer

Among the differentiated follicular cell-derived tumors, two major morphotypes are recognized: papillary thyroid carcinoma (PTC) and follicular thyroid carcinoma (FTC). The more numerous PTCs (70-90% of follicular cell-derived cancers) are subdivided into 12 subgroups, based on histological criteria, with the groups of classical PTC and follicular variant PTC accounting for the majority of cases. FTCs (10-30% of follicular cell-derived cancers) are subdivided into four subgroups, with the majority falling into the minimally invasive category. The prognosis for most patients within these common subgroups of PTC and FTC is usually excellent, with 10 year cause specific survival in excess of 95% (1-7). However, 15-40% will suffer a recurrence, mostly loco-regional, during their life time, sometimes several decades after the initial therapy (8, 9). It is this group of patients, who require prolonged follow-up, because delayed diagnosis and treatment of recurrence results in significant morbidity, in part related to damage to local structures by recurrent tumor, and in part due to problems related to the more extensive treatments required for bulky recurrence as compared to small recurrent lesions that are detected early. Tumor recurrence can be detected by a variety of means – examination of the neck, neck ultrasound, CT of neck or chest, or other body parts, radioactive iodine imaging, etc.; however, tumor markers, often provide a particular cost effective strategy to select patients for the (mostly) more costly definitive diagnostic procedures listed above.

Figure 1. Yearly incidence rates of thyroid cancer in the USA during 1974 through 2012.

Moreover, the importance of sensitive and (relatively) specific tumor markers for initial selection of thyroid cancer patients, who might need further work-up, becomes more important every year. Thyroid cancer incidence rates have increased exponentially around the world during the last 20-30 years (Figure 1) (1,3,5,6). At the same time, mortality rates have continued to decline. Since most other cancers with equal or higher incidence rates than thyroid cancer have much higher cause-specific mortality, thyroid cancer will likely be the 3rd most common tumor diagnosis in a living cancer patient by the end of the decade, just behind breast- and prostate cancer and ahead of colon cancer. Since tumor recurrence rates have remained much unchanged, this will result in a huge, and ever increasing, burden of follow-up.
Finally, as cause-specific survival continues to improve further, more and more patients might undergo more limited initial surgical procedures, reducing the diagnostic value of the current chief tumor marker in thyroid cancer, thyroglobulin (Tg), as we will discuss below.

**Thyroglobulin – the key thyroid tumor marker – strengths and limitations**

Tg is the chief protein produced by the follicular thyroid epithelium. It is the largest human protein with the mature dimer weighing ~660 kDa, about 10% of which are iodine and carbohydrates. It serves as scaffold and substrate for thyroid hormone synthesis, and subsequently functions as an intra-thyroidal storage reservoir for thyroid hormone (Figure 2). The chemistry involved in thyroid hormone synthesis involves iodination of tyrosine residues within Tg, followed by intramolecular coupling of certain mono-iodinated, or di-iodinated, tyrosines to form triiodothyronine (T3) or thyroxine (T4) (10). The entire process involves highly cell-toxic redux chemistry, mediated by enzyme complexes on the extracellular brush border of the apical (follicular lumen facing) membrane of follicular cells. The mature Tg thyroid pre-hormone is then stockpiled within the follicular space, forming the colloid. Colloidal Tg forms very large polymers and is poorly soluble (11). It serves mainly as long-term storage reservoir. Most of the thyroid hormone is not derived from colloid under normal conditions, but from recently synthesized and iodinated Tg, which is re-absorbed and proteolyzed, liberating T3 and T4, which are secreted via the basal membrane into the blood stream.

The relatively low disease-risk in patients with Tg levels between 0.2 and 2 ng/mL is most likely attributable to the presence of a small normal thyroid remnant in most of these patients. This remnant will secrete some Tg, as indicated earlier. To distinguish these individuals from the much smaller group of patients with residual or recurrent disease, it can be helpful to repeat the serum Tg measurement under conditions of thyrotropin (TSH) stimulation. This can be accomplished either through thyroid hormone withdrawal, or via administration of recombinant human TSH (rh-TSH). A rise of serum Tg concentrations above 2 ng/mL under stimulated conditions is usually considered an indication for further diagnostic work-up (21,22,24,28).

Remnant thyroid tissue is therefore one of the key limitations of the diagnostic usefulness of Tg. With the gradually increasing trend towards lesser surgical procedures, this limitation of Tg will become increasingly relevant. Currently, the European and North American Thyroid Cancer management guidelines still recommend total thyroidectomy (plus in many cases RAI remnant ablation) for the majority of patients (21,22,28). However, in Japan only ~5% of patients with tumors of 1-2cm size are treated by total thyroidectomy, whereas the reminder undergoes lesser procedures, most commonly lobectomy (61%). For even smaller tumors, which are nonetheless predominately (>50%) treated by total thyroidectomy in Europe or the USA, 40-73% of Japanese patients are stratified into watchful-waiting, i.e. no intervention and follow-up only (29,30), similar to what is now becoming popular in many Western countries for prostate cancer. Sooner or later similar approaches will gain ground in Australasia, Europe and the USA, and serum Tg measurements in their present form - and with the decision levels listed above - will become increasingly irrelevant.

Figure 2. Photomicrograph of a normal human thyroid gland (H&E stain), showing a transverse cut through follicles, filled with pink material, which is colloid, the storage form of thyroglobulin (Tg). The cartoon insert gives an enlarged schematic view of a single thyrocyte and outlines Tg and thyroid hormone synthesis and metabolism within the thyroid cell.

**Figure 2.** Photomicrograph of a normal human thyroid gland (H&E stain), showing a transverse cut through follicles, filled with pink material, which is colloid, the storage form of thyroglobulin (Tg). The cartoon insert gives an enlarged schematic view of a single thyrocyte and outlines Tg and thyroid hormone synthesis and metabolism within the thyroid cell.

By contrast, once the thyroid gland has been removed during initial cancer treatment, Tg is an excellent tumor marker for detecting potential residual or recurrent disease. Many studies have shown that it outperforms the traditional stalwart of thyroid cancer follow-up, radioactive iodine (RAI) whole body scanning, in such athyreotic patients (15). Following thyroid gland removal, serum Tg should become undetectable within several weeks, or a few months, depending on the preparative serum Tg concentration and the Tg clearance rate in a given patient. The latter can vary significantly between individuals. It appears to be biphasic, with a fast component of around 4-6 h half-life (T1/2) during the first ~48 h, which might reflect selective fast clearance of circulating monomers, partial monomers, and, perhaps, recently synthesized Tg dimers (16,17). The second component is slower, between 30-72 h T1/2, and might relate to clearance of stable dimers, or clearance of poorly soluble polycrystalline Tg that has been liberated from follicles during the surgical procedure (10,11,16,18-20). On average, Tg should fall in most patients (on thyroxine replacement) to below 0.1 or 0.2 ng/mL within 2-6 months after total thyroid ablation (total thyroidectomy +/- RAI remnant ablation, or near total thyroidectomy plus RAI remnant ablation).

It should be noted that the 0.1/0.2 ng/mL cut-off has a high negative predictive value (NPV), with a chance of only ~1% that such patients might have residual or recurrent disease. However, the corollary, that slightly higher serum Tg concentrations have a strong positive predictive value (PPV) for persistent/recurrent disease, is not true. Rather, there is a gradual increase in recurrence risk with increasing Tg levels, to about 2-5% at ~1ng/mL, 3-6% at ~1ng/mL, and 12-25% for the range from 2-10 ng/mL. Above 10 ng/mL, the likelihood of persistent/recurrent disease becomes quite substantial, 25-90% (15, 21-27).

There are a number of other factors, which reduce the utility of serum Tg measurements even in the current medical environment of total thyroid gland removal for thyroid cancer, namely antibody interferences, chiefly due to anti-Tg autoantibodies (TgAB) and, to a lesser extent, heterophilic antibodies (HAB). Both of these interferences are particularly troublesome in immunometric sandwich assays, which dominate Tg testing in the 21st century because of their high analytical sensitivity. TgAB or HAB may also cause problems in competitive Tg assays, but those seem to occur at a lesser frequency.
Detectable TgAB are found in 15-30% of thyroid cancer patients, at least 3 times the rate that is observed in the general population. In part this increase may be due to the fact that much lesser serum concentrations of TgAB are considered relevant with regards to the possibility of assay interference than would be considered indicative of autoimmune thyroid disease. Hence, published population surveys may not have counted low concentrations of TgAB as positive. In addition, it is known that tumor patients in general have higher rates of autoantibodies than the general population, often directed towards antigens that might be produced by their tumor (15). There are also suggestions that autoimmune thyroid disease might predispose to thyroid cancer (31). Regardless of the cause for TgAB, these autoantibodies gradually disappear in cured thyroid cancer patient over a period of several months (or a few years) following the initial thyroid removal (14,15). By contrast, patients with persistent/recurrent disease do not seem to clear their TgAB (14,15).

With regards to any patient with TgAB, immunometric Tg assays are likely to give false low serum Tg measurement results, presumably due to interference of TgAB with Tg capture, detection, or both. The frequency and dose-dependency (of both TgAB and Tg) of this interference depends on the Tg and TgAB assays used, and also differs between patients (15,32). Competitive assays do not seem to be affected as frequently, perhaps because the polyclonal antibody preparations that are typically used in a competitive assay have a wider epitope spread. Unfortunately, almost all competitive immunooassay have insufficient analytical sensitivity to satisfy current practice guidelines, with most of them struggling to detect Tg concentrations that are <3-5 ng/mL. If interference does occur in a competitive immunoassay, it will mostly result in a false high result, which can, of course, also affect optimal patient management, just in a different way.

HABs are antibodies in a patient’s serum that can cross-react with the antibodies that are used in an immunoassay. In an immunometric sandwich assay HABs might bind simultaneously to both capture and detection antibody, thus simulating the presence of analyte in its absence, or, if some analyte is also present, leading to a false high measurement. HAB are found in 3-6% of the general population, and assay manufacturers therefore add ‘blocking reagents’, usually non-specific IgG from the same species as the capture and detection antibodies, to immunometric assays. This should cut the frequency of HAB interferences to a fraction of a percent. Unfortunately, when very high analytical assay sensitivity is required, such as in high sensitivity Tg assays, adding too much blocker can compromise assay performance. In addition, tumor patients tend to have higher rates of HAB positivity than the general population. Consequently, immunometric Tg assays might be affected in 0.1-3% of patients (33). Again, competitive assays are affected less frequently.

Between TgAB and HAB, anywhere from 15% to over 30% of serum Tg measurements that are performed with immunometric sandwich assays are compromised. Moreover, while it is possible to identify most, but not all, samples that might be affected (easier for TgAB than for HAB), it is almost always impossible to determine the “true” serum Tg concentration in these patients. Serial measurement of TgAB concentrations might in some patients be a suitable surrogate for accurate Tg measurements (31), but any change in Tg or TgAB assay during this time will void such efforts, principally because of the poor quantitative agreement between different TgAB assays (32,34).

Potential solutions for the limitations of Tg testing

The antibody related problems of current Tg immunoassays have been the major driver to find alternative tumor markers to Tg, or different Tg measurement methodologies. The quest for alternative tumor markers started during the 2nd half of the 1990s and has centered on various molecular markers, while more recent alternative Tg measurement techniques have focused on mass spectrometry.

Molecular thyroid cancer tumor markers

Molecular thyroid tumor markers fall, with some overlap, into two categories: organ specific markers, similar to Tg, and tumor specific markers.

Organ-specific circulating molecular tumor markers

The evaluation of organ-specific molecular markers has focused on detecting messenger RNA (mRNA) for thyroid specific proteins in patient blood samples. Investigators have developed assays for the detection of circulating mRNA species coding for thyroid peroxidase (TPO), the sodium iodine symporter (NIS), pendrin (PDS), the thyrotropin receptor (TSHR) and Tg. Of these, TSHR and Tg mRNA measurements have received the most attention, with TSHR mRNA measurements being the only test that is currently used clinically (on a relatively small scale). However, Tg mRNA was the 1st marker that was investigated, and it remains the most studied. In 1996 Ditkoff et al. published a paper in which they stimulated thyroid cancer cell lines with recombinant human TSH (rh-TSH) and then detected Tg mRNA transcripts in patients’ blood samples, with an analytical sensitivity equivalent to the Tg mRNA content of about 200 tumor cells per mL of blood. (35). Since Tg protein expression is highly restricted to the thyroid gland, they surmised that Tg mRNA transcription would likewise be restricted to thyrocytes, and, hence, any Tg mRNA detected in circulation would have to come from circulating normal or malignant thyrocytes.

When they applied their assay to individuals with benign thyroid diseases and healthy controls, they detected no amplification product. By contrast, in the blood of postoperative thyroid cancer patients they detected Tg transcripts in all 9 patients with persistent/recurrent disease and in 7 of 78 patients who had no evidence of disease-persistence. This translated into cancer follow-up test sensitivity and specificity of 100% and 91%, respectively, and, based on the about 10% prevalence of patients with active disease in the study, in respective positive and negative predictive values of ~56% and 100%. These figures were as good, or better, than what could be achieved with the serum Tg assays available around that time (detection limits 2-5 ng/mL). Understandably, these results caused quite a stir in the thyroidology community, in particular, as one could reasonably expect that TgAB, which are directed against the Tg protein, but not against the Tg mRNA, would most likely not interfere in this method. Not surprisingly, several other groups subsequently developed their own Tg mRNA assays (35-60). While most of these focused on thyroid cancer follow-up and addressed very similar questions to the original paper in roughly comparable patient populations, they all varied from the original paper, and from each other, in many methodological aspects, including criteria for disease recurrence, stratification of patients into different groups, inclusion or not of various control populations, sample collection procedures, RNA extraction protocols, rt-PCR methods, primers and probes, and whether the assays were qualitative or quantitative (and cut-offs for the latter) and what their detection limits were. Different studies also varied as to whether they included a comparison with serum Tg measurements, as well as with regards to which serum Tg assay was used, what serum Tg assay cut-offs were employed and whether the measurements were performed with the patients on thyroxine or after thyroxine withdrawal/rh-TSH stimulation. All studies were essentially retrospective and most groups published only a single paper, but no further validation/confirmation study. Those, who performed more than one study, did not make it clear if, or how much, overlap there was in the patient populations that were used in the different studies. The result of all this was that some studies seemed to confirm the initial data, while others showed little diagnostic value in Tg mRNA measurements, as is summarized in Table 1.
If one uses the increment in post-test PPV and NPV above the corresponding pre-test values as a measure of diagnostic utility of a test, the combined results of the 21 published studies (8 of which included two or more subgroup analysis or different types of data analysis) show an average 1.86 fold increase in PPV (range 0.57 fold - worse than no testing – to 9 fold - substantial improvement through testing), while the corresponding figures for NPV range from 0.69 fold to 2.1 fold (average 1.21 fold).

With the high degree of design variability between the various studies it is hard to pinpoint the causes for these discrepant results. Given such variable study outcomes, and considering the instability of mRNA as an analyte, with all the associated sample collection, processing and storage issues/problems, the interest in Tg mRNA measurements started to decline by the middle of the last decade, and to my knowledge there are no clinical assays in routine use today.

TPO, NIS and PDS mRNA have fared even worse. They were studied less extensively, usually in conjunction with Tg mRNA, and there have been very few studies, all with inconclusive results. However, TSHR mRNA testing has arguably eclipsed Tg mRNA. It is regularly used at the Cleveland Clinic, where it was pioneered, and can be ordered by outside physicians through the referral branch of the Cleveland Clinic Laboratories (although, at this time of writing, testing has been suspended due to problems obtaining control materials). The Cleveland Clinic group has published a number of articles on TSHR mRNA testing in blood, which showed quite favorable results for a range of applications, which have included (i) the typical application, detection of residual/recurrent disease in treated patients (Tg mRNA was included in one study) (52,61), (ii) as an adjunct to cytopathology in the initial diagnosis of thyroid cancer (62,63), and (iii) risk stratification in patients with papillary microcarcinoma (61). Unfortunately, it is not easy to determine the proportion of the patients who were included in only one as opposed to several of these studies. In addition, a large number of subgroup analyses were performed. Consequently, it is a little challenging to come to any firm conclusions about the actual clinical performance of the assay.

The most useful paper for this purpose is probably a summary of the group’s overall experience published in 2010 (64), which gives a good aggregate view of the various applications of the testing and the results obtained. Interestingly, the most frequent application seems to be use of the blood TSHR mRNA test as an adjunct in the initial diagnosis of thyroid cancer. TSHR blood mRNA measurements in these patients yielded consistently high PPVs of 81-100%, depending on the subgroup studied; however, NPVs were less compelling, ranging from 39-84%. When combined with accepted ultrasonic criteria to determine malignancy in thyroid nodules, the NPV rose to 95% (PPV stayed at 88%), a threshold that the USA National Comprehensive Cancer Network (NCCN) thyroid cancer guidelines recognize as being acceptable for deferral of surgical intervention (24). For the more conventional application of disease surveillance during follow-up, the same publication reports PPVs and NPVs of 83% and 88%, respectively, in TgAB negative patients. The corresponding figures for TgAB positive patients are 100% and 96%, but this group contained only 4 (of 60) individuals who were TSHR mRNA positive, suggesting that these PPV and NPV calculations will have wide confidence intervals. With regards to a comparison with serum Tg measurements one has to go back to an earlier publication (65), which shows diagnostic sensitivities and specificities for TSHR mRNA in blood of 50%-66% and 77-88%, respectively, while the comparative figures for serum Tg measurements are 55-94% and 96%. The "either – or" combination of both tests, results in 90% sensitivity and 94% specificity. It is interesting to note, that the cohort of patients, in whom TSHR mRNA measurements seem to have an edge over serum Tg, are those who are TgAB-positive.

Overall, it appears that TSHR mRNA measurements might be useful in some patients with TgAB, specifically if Tg is <2 ng/mL; in individuals with higher levels, further diagnostic work up is indicated anyway, and TSHR mRNA contributes little.
The absence of TSHR mRNA in healthy controls also moves this marker from the purely organ specific category into the potentially tumor specific realm. Further work in this area should be encouraged.

Finally, it would be reassuring if a larger number of other groups would independently reproduce some of the Cleveland Clinic group’s work. I am only aware of three publications by other groups that have evaluated TSHR mRNA as a tumor marker in thyroid cancer (all for the purpose of follow-up). Of these, one showed a diagnostic performance on T4 and off T4 comparable to Tg mRNA, and slightly worse than serum Tg (51), while the other two failed to show any clinical utility (60,66).

**Tumor-specific circulating molecular tumor markers**

This category contains two groups of markers, (i) molecular/genetic oncofetal markers that might be expressed by many different types of human neoplasms, much akin to established oncofetal protein markers (e.g. CEA), and (ii) mutated/rearranged oncogenes/tumor suppressor genes, which are very frequently found in thyroid carcinomas; in some instances they might be modestly to strongly specific for thyroid tumors, but mostly they are also present at low to medium (and sometimes high) rates in other human tumor types. Notably, all of these potential markers in the second group are DNA-based, a much more stable and easier to work with sample type than mRNA.

With regards to the molecular oncofetal markers, the only one that has been evaluated to date in follicular cell-derived thyroid carcinoma is oncofetal fibronectin (onfFN) mRNA. A proof of principle publication in 2005 showed that onfFN transcripts could be detected in the blood of 6 of 9 patients with metastatic thyroid carcinoma, but not in 8 control subjects (66). The same group subsequently expanded their studies to 25 control subjects, and 95 thyroid cancer patients. Amongst the latter, 64 were disease free, 23 had local residual/recurrent disease, and 8 distant metastases. They found significantly higher onfFN mRNA concentrations in the patients with active disease; however, there was significant overlap between the ranges, with many disease-free individuals having high onfFN mRNA levels and a significant number of patients with active disease having relatively low levels (67). Two subsequent publications by other investigators have since confirmed the large diagnostic overlap, essentially disqualifying onfFN mRNA as a useful thyroid cancer tumor marker (68,69).

The second group of tumor-specific markers includes mutated HRAS, NRAS and KRAS genes (codon 12, 13 or 61 mutations), which are collectively found in 8-20% of both PTC and FTC, mutated BRAF (almost exclusively c.1799T>A, p.V600E; 40-80% of PTC) and at least 15 variants of oncogenic rearrangements involving the RET proto-oncogene (RET/PTC1, collectively <5%-20% of PTC) and various fusion partners (RET/PTC1 is by far the most common, followed by RET/PTC3, the rest are rare), which are exclusively seen in PTC, and the family of PAX8/PPARγ rearrangements (at least 4 variant transcripts), which are confined to FTC (30-60% of FTC) (70-74). Of these markers, KRAS mutations are also found at high frequency (>20%) in various other human cancers (e.g. colon, pancreas and lung), while BRAF is seen in most melanomas (>80%) and many other cancers (at rates of 5-20%) (72,75,76). RAS and BRAF mutations are therefore more tumor-specific than organ-specific. By contrast, RET/PTC1 and PAX8/PPARγ rearrangements are seen exclusively in FTC and FTC, respectively, and can, hence, be considered both organ and tumor specific. However, an additional fact to note is that the term “tumor-specific” includes adenomas and carcinomas for all the above markers, with the exception of BRAF, which appears specific for malignancy in the case of thyroid tumors.

The most studied of these potential markers is BRAF, a tempting target because of its high prevalence in the most common thyroid cancer morphotype, PTC. Moreover, nearly all disease causing mutations involve a single nucleotide change – c.1799T>A, p.V600E – greatly facilitating assay design. There have been five studies that measured mutant BRAF in circulation in thyroid cancer patients (77-81). Of these, two can be considered “proof-of-principle” studies (77,79), while the other three attempted to establish the clinical performance of mutant BRAF detection assays (78,80,81).

One of these three studies was performed by our group and focused on thyroid cancer follow up. Whole blood samples were obtained prospectively from 193 patients with thyroid cancer (173 with PTC, and 20 with FTC or MTC) during routine follow-up visits. DNA was extracted from the blood samples and subjected to semi-quantitative real-time PCR with allele specific primers for wild type BRAF and mutant BRAF (c.1799T>A). We detected circulating mutant BRAF in 20 of 173 PTC patients (11.6%) and none of the non-PTC patients. The presence of mutant BRAF in blood correlated with disease status (relative risk of recurrence 2.55), but in this cohort of patients, who almost all had undergone complete thyroid removal, the testing did not show any clear advantage over the combination of unstimulated and stimulated serum Tg measurement.

The other two studies focused on detection of circulating mutant BRAF for initial diagnosis of thyroid cancer. Rather than whole blood, serum (80) or plasma (81) samples, respectively, were collected prospectively from patients presenting for initial thyroid cancer diagnosis. The study by Kwak et al. (80) included only serum samples from patients with later surgically confirmed PTC, while the paper by Pupilli et al. (81) included plasma samples from subjects with biopsies that were non-diagnostic, benign, follicular lesions, suspicious for malignancy and malignant, as well as plasma samples from healthy volunteers and individuals with non-nodular thyroid disorders. Both groups used real-time PCR; the Pupilli et al. group used allele specific priming, while the Kwak paper used a “mutation specific” commercial real-time PCR kit – it remains unclear from the paper, whether the kit employs allele specific amplification or allele specific detection. The results of the two studies differed substantially. The Korean group failed to detect amplifiable mutant BRAF in any of their samples, while the Italians found mutant BRAF in all subject groups, including the normal controls. This suggests that their allele specific amplification was not as selective as would be desirable. However, there was a steady increase in the proportion of detected mutant alleles versus wild type alleles across the subject groups from <2% in the healthy controls to about 27% in the suspected or confirmed cancer groups, with the other categories falling in between. Likewise, there was a significant difference in relative mutant allele concentrations between patients who were confirmed to have PTC at surgery versus those, who proved to have benign lesions. Furthermore, most patients with PTC showed a decline in the proportion of mutant alleles after tumor removal. Finally, the group also evaluated the positive and negative predictive value for predicting final diagnosis in the group of patients with follicular lesions, using a cut-off of 2.65% mutant alleles. The resultant PPV and NPV were 33% and 80%, respectively.

In summary, it appears that circulating mutant BRAF can be detected in some PTC patients. However, there are insufficient data to allow any reliable conclusions about clinical usefulness. In addition, there is a lack of standardization of sample type and assays, which further confounds the issue.

Other tumor-specific markers have fared even worse. The only other tumor-specific marker that has been studied is RET/PTC1, which was evaluated in a single study in 1998 (37). The RET/PTC1 fusion transcript was detected by rt-PCR in the blood of 1 of 24 patients with thyroid cancer and in none of 20 control patients (goiter and adenoma). The small number of patients in this study and the relatively low prevalence of RET/PTC1 in the tested patient population preclude any useful conclusions from this observation.
In conclusion, neither the organ-specific-, nor the tumor-specific circulating molecular tumor markers have thus far lived up to expectations. There are some tantalizing glimpses of mainstream utility, but at this stage none of these markers can replace serum Tg measurements or address its shortcomings consistently.

Alternative Tg measurements methods to address Tg immunoassay shortcomings
In parallel with the quest for molecular thyroid tumor markers there have been increasing efforts to develop new serum Tg assays that are not affected by the presence of TgAB. The first assay that laid claim to this was the SANOFI-Pasteur immunoradiometric Tg assay. In the initial publication, this assay, which used several different capture antibodies, seemed to be largely immune to TgAB interferences (82). However, it was quickly shown, that, while the multi-antibody capture approach reduced the TgAB problem, it did not abolish it (83). With extensive use in a modified incarnation, as the Beckman Access automated thyroglobulin immunoassay, these initial observation have been corroborated; the assay appears to have lesser rates and magnitude of interference compared with many others, but it is by no means free of these problems (15,31,32).

More recently, efforts have been directed towards the development of mass spectrometry (M) based assays of Tg. Protein identification and measurement by M has a long tradition. The M workflow for protein measurements usually involves a digestion of the sample with trypsin. Trypsin cleaves proteins in a predictable fashion into peptides, which can be measured by M and identified by protein database matching. These qualitative methods can be adapted for clinical quantitative measurements. In the case of Tg, the tryptic digest will cleave all proteins in the sample, including Tg and any TgA that might be present. One can then specifically look for (one or several) tryptic peptides that are prototypic for Tg (based on predicted cleavage), without any interference by TgA. Quantification is achieved by including a standard curve based on synthetic versions of the measured peptide(s), and adding non-radioactive isotopic version(s) of the peptide(s) in defined concentration(s) as internal standard(s) to the sample(s) before analysis. In practice, things are a little more difficult. One faces a formidable signal to noise problems in identifying the quite low concentration Tg peptides in a trillion-fold higher background abundance of all the other peptides from all the other proteins. To overcome this problem, some form of sample enrichment is required. Separation of proteins based on size (Tg, TgAB, and Tg-TgAB complexes are very large proteins) can be used before trypsin digestion. Even more effective is immune affinity purification of the desired Tg target peptide(s) from all the other peptides. A good antibody against the target peptide can achieve dramatic enrichment. Using one or both of these enrichment techniques, Tg measurement by M becomes possible. By limiting oneself to one or two peptide targets that ionize well, optimizing chromatography, and using selective reaction monitoring on high sensitivity mass filtering tandem mass spectrometers, rather than the higher resolving, but (typically) less sensitive, high resolution instruments used in proteomics research, analytical sensitivities can be achieved that are comparable to modern immunometric Tg immunoassays. Figure 3 shows an example of the typical overall workflow.

Several groups have now published their experience with such approaches (84-86). Analytical sensitivity for the earliest assay was around 4 ng/mL (84), but has improved with more recent iterations to about 0.5 ng/mL (85,86), comparable to most commercial immunoassays and just a little higher than the EIASON and Beckman ultrasensitive Tg immunoassays.

There is typically good agreement with the results obtained by immunoassays in TgAB negative patient samples (Figure 4).

Figure 3. Schematic depiction of the workflow for current liquid chromatography, tandem mass spectrometry (LC-MS/MS; HPLC-MS/MS) measurement of Tg.
Figure 4. Method comparison of Tg measurement by LC-MS/MS with automated immunoassay (IA) measurement (Beckman assay) in anti-thyroglobulin autoantibody (TgAB) negative patient samples with Tg concentrations of >1 ng/mL by Beckman IA. There is a very tight correlation between the two assays. However, the LC-MS/MS assay shows approximately a 20% downward bias. This is most likely due to mis-calibration of the Beckman assay, which becomes apparent when international standard material (CRM457) is run as “unknown” in the Beckman assay (insert figure).

Figure 5. Method comparison of Tg measurement by LC-MS/MS with automated immunoassay (IA) measurement (Beckman assay) in TgAB positive patient samples with Tg concentrations of >1 ng/mL by Beckman IA. There is still a fairly tight correlation between the two assays. However, the LC-MS/MS assay now gives higher results; the bias has changed from ~0.8x to ~1.3x compared with the method comparison in TgAB negative patients. This is most likely attributable to approximately 60% systematic under-recovery in the Beckman assay measurements of these samples due to false low TgAB interference.
Spike and mixing experiments of exogenous Tg into TgAB positive samples confirm that the Tg MS assays are not affected by the presence of TgAB, while immunoassays, predictably, under-recover. The absence of TgAB interference in Tg MS measurements is further supported by comparisons with Tg immunosassays in TgAB positive samples. Here, the Tg immunosassays shows a low bias compared to the Tg MS assays (Figure 5), and some of the TgAB positive samples have undetectable Tg, but measurable amounts are found by the MS assay(s) (Figure 6). It is of note, though, that at least in the Beckman Tg immunosassay, the Tg concentrations that are found by Tg MS in samples with undetectable Tg concentrations by immunoassay (<0.1 ng/mL) are not dramatically high. Whether this holds true for all Tg immunosassays is currently unknown; the laboratories that have developed Tg MS assays all use the Beckman Tg immunosassay. It is possible that this is could be due to the Beckman assay being relatively resilient to TgAB interferences. However, it is also possible that TgAB cause accelerated clearance of Tg; after all this is one of the key biological function of immune complexes. If this was the case, then Tg MS assays might not improve the diagnostic performance in TgAB positive patients with undetectable serum Tg as much as one might have anticipated. Clearly, comparison studies of Tg MS assays with multiple Tg immunosassays in reasonably large patient cohorts with known clinical outcome are required to address these uncertainties.

With regards to HAB interferences, there are no published data for Tg MS assays. One would, however, expect that Tg MS solves these problems, and in our laboratory we have anecdotal evidence that this indeed so.

Summing up, Tg measurement of serum Tg addresses the degraded clinical diagnostic performance of Tg immunoassay in TgAB positive patients and in HAB positive patients at least in part. Whether the Tg MS assays represent a complete solution must await further studies that correlate outcome data with assay data.

**Future requirements for circulating thyroid cancer markers**

While the new Tg MS assays seem to address, at least in part, perhaps substantially, the antibody interference problems of Tg immunosassays, they do not solve the potential future problems related to the anticipated increase in the numbers of patients with significant amounts of residual thyroid tissue following primary treatment.

Tumor-specific (or partially tumor-specific) markers should in theory be able to fill this gap. However, as we have seen, despite almost 20 years of work on these markers, they have thus far not fulfilled their promise. A new avenue that is being pursued is measurement of circulating miRNA. Micro RNA expression patterns change in a wide range of malignancies (and many other diseases), and significant quantities of potentially disease-associated miRNAs can be found in the circulation (87-90). Most of these miRNAs are quite stable, unlike mRNA, making them a more attractive analytical target for clinical applications. Unfortunately, thus far, clinical applications have not gone much beyond the proof of principle stage, or have not been independently verified. In part this might be the case, because many of the pre-analytical variables that affect the concentrations of various circulating miRNAs are not yet fully understood (91,92).
Massively parallel sequencing of circulating cell free DNA is another possibility. This methodology has found widespread international use in clinical antenatal fetal aneuploidy detection and has proven reproducible and reliable for this purpose (93-96). One could envisage that such an approach could be adapted to cancer follow-up, perhaps even individualized to the specific somatic genetic changes that were in a given patient’s primary tumor. The limiting factors could be cost – which is likely to decrease, but perhaps not sufficiently – and analytical sensitivity. The tumor DNA would presumably be a small fraction of all circulating cell free DNA, much like fetal DNA concentrations in maternal blood rarely exceed 20% of all cell free DNA; at less than 10% fetal DNA the antenatal aneuploidy assay become unreliable (93-96). From the BRAF studies we know that in thyroid cancer patients circulating mutant DNA is mostly below 1%, which would suggest that massively parallel sequencing may not be a viable solution to this problem. Of course, the field of next generation sequencing is evolving so rapidly that this could change in the near future.

In either case, there may be some non-molecular (conventional) tumor marker approaches that might allow detection of recurrence in patients with sizable thyroid remnants. These approaches center again on Tg, or more accurately on distinguishing tumor-derived Tg from Tg that is secreted by normal remnant thyroid tissue. It has long been known that tumor-derived Tg lacks iodination, is poorly glycosylated and shows reduced sulfation of chondroitin sulfate (97-102). These differences may allow specific detection of tumor Tg within a substantial background of normal Tg. However, currently there are not such assays available that have been developed beyond the proof of principle stage.

Finally, in the absence of new modified conventional or molecular tumor markers, the thyroid remnant problem might still be in part solvable by generating much more complete, comprehensive and detailed reference ranges for Tg, which correlate remnant size and TSH levels in large numbers of subjects with serum Tg concentrations. This would allow placing individual patients into bands of expected Tg concentrations. If this is combined with much better assay standardization, or, of course, if a single type of assay is used on a given patient throughout his or her life, such an approach should still allow relatively early detection of recurrence.

Acknowledgements

- My colleagues Ravinder J Singh, PhD and Alicia Algeciras-Schimnich, PhD, who have been involved in most aspects of planning and execution of molecular testing efforts and the development of thyroglobulin mass spec testing methods.
- The development scientists/technologists, who have driven the development of molecular and mass spec tests that are mentioned in this review: Kendall W Cradic, MS; Dragana Milosevic, MS; Brian C Netzel, MS; and Robert Taylor, BS.
- Various clinical colleagues, chiefly, Bryan McIver, MB BS PhD.
- Norman L Eberhardt, PhD; Honey Reddi, PhD, for supporting related basic science activities in their laboratory, and David R Barmidge, PhD for supporting all proteomic efforts in the clinical proteomics core facility.

Author information

Stefan KG Grebe, MD PhD FRACP DABCC, Professor of Laboratory Medicine and Pathology

Department of Laboratory Medicine and Pathology, Mayo Clinic Rochester, 200 1st Street SW, Rochester, MN 55905, USA.

Email: grebe.stefan@mayo.edu

References

8. Leung AM, Dave S, Lee SL, Campion FX, Garber JR, Pearce EN. Factors determining the persistence or recurrence of well-differentiated thyroid cancer treated by thyroidectomy and/or radioiodine in the Boston, Massachusetts area: A retrospective chart review. Thyroid Res 2011; 4: 9.


32. Spencer C, Petrovic I, Fatemi S. Current thyroglobulin autoantibody (TgAb) assays often fail to detect interfering TgAb that can result in the reporting of falsely low/undetectable serum Tg IMA values for patients with differentiated thyroid cancer. *J Clin Endocrinol Metab* 2011; 96: 1283-1291.


45. Eszlinger M, Neumann S, Otto L, Paschke R. Thyroglobulin mRNA quantification in the peripheral blood is not a reliable marker for the follow-up of patients with differentiated thyroid cancer. *Eur J Endocrinol* 2002; 147: 575-582.


67. Wehmeier M, Petrich T, Brand K, Lichtinanhagen R, Hesse E. Oncocytotelial fibroblastic mRNA is highly abundant in the blood of patients with papillary thyroid carcinoma and correlates with high-serum thyroid-stimulating hormone levels. Thyroid 2010; 20: 607-613.


71. Iaczewski KA, Reddi HV, Grebre SK, Eberhardt NL, McVier B. The role of the PAX8/PPARgamma fusion oncogene in thyroid cancer. PPAR Res 2008; 2008:672829.


78. Cradic KW, Milosevic D, Rosenberg AM, Erickson LA, McVier B, Grebe SK. Mutant BRAF(T1799A) can be detected in the blood of papillary thyroid carcinoma patients and correlates with disease status. J Clin Endocrinol Metab 2009; 94: 5001-5009.


NZ J Med Lab Science 2013


