

GUIDELINES FOR FORMULARY EVALUATIONS

Nullius in verba



GUIDELINES FOR FORMULARY EVALUATIONS

[PROPOSED]

PROGRAM IN SOCIAL AND ADMINISTRATIVE PHARMACY

COLLEGE OF PHARMACY

UNIVERSITY OF MINNESOTA

Paul C Langley Ph.D., Adjunct Professor

Version 2.0: December 2016

For further information on implementing these guidelines please contact:

GFE, College of Pharmacy, University of Minnesota

308 SE Harvard Street, Minneapolis MN 55455

Tel: (612) 624-1000

Email: schom010@umn.edu



GUIDELINES FOR FORMULARY EVALUATIONS

NEXT GENERATION SEQUENCING

THE DECEMBER 2016 REVISION

The December 2016 revised version (2.0) of the Minnesota *Guidelines for Formulary Evaluation* has been prompted by the growing interest and, to a limited extent, introduction and coverage of Next Generation Sequencing (NGS) platforms. The introduction NGS platforms promise a major rethink in how formulary committees are to evaluate claims, not only for new pharmaceutical products, but for submissions to reimburse individual NGS platforms. At the same time, while the acceptance of NGS platforms as an integral part of treatment guidelines and as an arbiter of therapy choices promises a major shift in the process of drug discovery and treatment practice, the critical evidentiary standards remain unchanged. All claims, whether they be for specific compounds, specific platforms or for the formal introduction of NGS platforms into therapy guideless must be robust, credible, evaluable and replicable. To achieve this formulary committees must insist on submissions and a commitment to evaluation that meet these standards of normal science.

A commitment to NGS in therapy choice presents formulary committees with two decisions: (i) the choice of NGS platform for approval and reimbursement to support therapies for target populations within a disease area and (ii) the adoption of NGS sequencing to support the approval and reimbursement of new drug products to complement therapy choices for target populations within a disease area.

GUIDELINES FOR FORMULARY EVALUATIONS

TABLE OF CONTENTS

	Page
1. BACKGROUND	
1.1 Credibility and Replication	6
1.2 Next Generation sequencing	7
1.2.1 NGS and Treatment Guidelines	9
1.2.2 NGS and Risk Stratification	10
1.2.3 Barriers to NGS Implementation	10
1.3 Acceptance of Modeled Claims	11
1.4 Categorizing Formulary Submissions	11
1.4.1 Standard Submissions	12
1.4.2 Single Molecule Submissions	13
1.4.3 NGS Platform Submissions	13
1.4.4 Therapy Target Submissions	14
1.5 Quality Adjusted Life Years	14
1.6 Submission Protocol	14
1.7 Target Populations	15
1.7.1 Drug and Platform Costs in Target Populations	15
1.8 Options in Claims Evaluations	16
1.8.1 Comparative Effectiveness Research	16
1.9 Value Claims	17
1.10 Adherence and Persistence	17
1.10.1 Physician Uptake	18
1.11 Comorbidities	19
1.12 The Evidence Base	20
1.12.1 Diagnostic Accuracy and Quality	20
1.12.2 Systematic Reviews	20
1.12.3 Reporting Randomized Trials	20
1.12.4 Evidence Hierarchy	20
1.12.5 Evidence Base in Oncology	21

GUIDELINES FOR FORMULARY EVALUATIONS

2. EVALUATING FORMULARY SUBMISSIONS: KEY ELEMENTS

2.1 Product and Comparator Descriptions and Pricing	22
2.2 Target Population and Place in Therapy	22
2.3 Significant Adverse Events and Contraindications	22
2.4 Primary and Secondary Outcomes	23
2.5 Direct and Indirect Comparisons	23
2.6 Replication	23
2.7 Adherence and Persistence	24
2.8 Completed Evaluations	24
2.9 Evidence Summaries	24

3. VALUE AND COST-OUTCOMES CLAIMS

3.1 Scope and Options Value and Modeled Cost-Outcomes Claims	25
3.2 Claims Evaluation Protocol	25
3.3 Value Claims and Model Structures	26
3.4 QALYS and Cost-Outcomes Thresholds	26
3.5 Adherence and Persistence in Modeled Claims	27
3.6 Dispensing with a Modeled Claim	27
3.7 Concurrent Submissions	27

4. PRODUCT AND COMPARATOR CLAIMS ASSESSMENT

PROTOCOL

4.1 Protocol Standards (PROST)	28
4.2 Format for Product Protocol Submission	28
4.3 Questions a Formulary Committee Should Ask	29

GUIDELINES FOR FORMULARY EVALUATIONS

5. NEXT GENERATION SEQUENCING PLATFORM ASSESSMENT PROTOCOL	
5.1 Comparative NGS Platform Assessment	31
5.2 Evidence Base for Comparative NGS Assessment	31
5.3 Questions a Formulary Committee Should Ask	32
6. REQUEST FOR SUBMISSIONS	34
REFERENCES	35

1. BACKGROUND

1.1 Credibility and Replication

If formulary decisions to admit or retain pharmaceutical products and devices are to be credible they must recognize the standards of modern science. The focus of the *Guidelines for Formulary Evaluations (GFE)* lies in the recognition that if claims are made to support pharmaceutical products and devices, they must be credible, evaluable and reproducible¹. Claims must be presented in a testable form that allows feedback to formulary committees as part of ongoing disease area and therapeutic reviews. This evidentiary standard applies equally to claims for comparative clinical effectiveness as well as to claims for cost-effectiveness. If non-evaluable claims are presented by manufacturers then they should either be reformulated or put to one side as lacking credibility for formulary decisions.

This standard for experimentation based on evaluable claims is unexceptional. It has been recognized as the core of the scientific method since the mid-17th century². Unfortunately current standards for formulary submission fail to recognize this standard. Claims are submitted to support formulary decisions that are incapable of falsification and replication^{3 4}. This is true, not only of the Academy of Managed Care Pharmacy (AMCP) *Format for Formulary Submissions (Version 4, April 2016)* but also of ‘gold standard’ reference case claims mandated by agencies such as the National Institute for Health and Care Excellence (NICE) in the UK, PHARMAC in New Zealand, the Health Information and Quality Authority (HIQA) in Ireland and the European Union.^{5 6 7 8 9 10 11 12 13} These same criticisms apply to modeled or simulated claims published in journals such as *Value in Health*, *PharmacoEconomics* and the *Journal of Medical Economics*^{14 15 16}.

If the NICE reference case standards are applied, for example, cost-utility models are submitted to formulary committees which make the comparative case for chronic disease interventions in lifetime discounted cost-per-QALY terms. These claims are, quite clearly, impossible to evaluate; nor were they ever intended to support evaluable outcomes¹⁷. Yet formulary committees are asked to accept such modeled or simulated claims even though they have no idea whether they are right or even if they are wrong. The argument that because a modeled or simulated claim demonstrates a sufficient correspondence to reality it should necessarily be accepted in the absence of experimentation is not a basis for informed formulary decision making. Such claims are nothing more than thought experiments or imaginary worlds and are not relevant to evidence based formulary decision making^{18 19}. The claims fail the test that demarcates science from pseudoscience²⁰.

Of particular note in the US market are claims presented, together with ‘notional discounts’ for pricing set against willingness to pay threshold by the Institute for Clinical and Economic Review (ICER). These claims, which mirror the reference case cost-per-QALY methodology utilized by NICE also typically fail the standards of credibility, evaluation and replication²¹. If these standards are applied then these claims should also be rejected unless, on reformulation, claims can be presented in a form that meets the standards of normal science. To base formulary decisions on clinical and cost-effectiveness claims that fail to meet the standards of normal science in both testability and replicability is an untenable situation. With increasing

GUIDELINES FOR FORMULARY EVALUATIONS

concerns over the ability to replicate claims from Phase 3 trials and output switching in reporting phase 3 results, formulary committees have to be aware of the need to establish the credibility of claims^{22 23 24}.

1.2 Next Generation Sequencing

Next generation sequencing (NGS) has the potential to disrupt not only the accepted process of drug development but also the hurdles a drug manufacturer would be expected to face in securing formulary approval and a possible premium price for a new compound²⁵. NGS (sometimes referred to as massive parallel sequencing) is a high throughput approach to DNS sequencing via a spatially separated, clonally amplified DNA templates or single DNS molecules. Example of NGS applications includes prenatal testing, identification of rare genetic variants and the efficient detection of germline or somatic mutations in cancer genes. In cancer, as the most obvious candidate for NGS, the accurate molecular subtyping of disease offers a path to the credible linking of molecularly targeted therapies to mutations or mutation clusters. At the same time tracking via NGS platforms potentially allow treating physicians to monitor cancer progression and adjust targeted therapies.

At the same time, the introduction of NGS targeting, gives added strength to the arguments for credibility, evaluation and replication in submitting formulary claims. The reason for this is that once the NGS paradigm is accepted, the choice of therapy is not restricted to a single molecular target that characterized the earlier attempts to match therapies to somatic mutations. NGS captures, through full genome sequencing, the range of mutations present in treating populations. The heterogeneity of tumor expression means that the NGS platform can identify, in a target population, the distribution of mutation load where the population is characterized by potential therapy pathways defined by mutations or mutation clusters.

Once NGS targeting is introduced the choice of therapy, in hard tumor oncology for example, relies upon an assessment of the mutation load and the distribution of mutations and mutation clusters across the target patient population. The NGS platform attempts to link therapies, both single and in combination, to individual mutations and mutation clusters. The matching of mutations to therapies where there is prior evidence of their efficacy in attacking specific mutations offers the prospect of improved outcomes and lower costs. While such claims are still subject to the standards of credibility, evaluation and replication, the focus of formulary assessment moves away from 'single compound' submissions to 'multiple compound' claims, where the case is made in terms of an 'across the board' multiple therapy intervention where each therapy pathway is tied to a mutation cluster.

The introduction of NGS sequencing may be characterized as one of creative destruction, where its adoption in personalized medicine sets in train a mechanism of ongoing product and process review. A mechanism driven by continuing modifications and extensions to NGS platforms as the understanding of the role of mutations and mutation load in therapy choice expands. At the same time this mechanism has significant implications for the continued revision of treatment guidelines and their adoption of NGS as integral parts of the treatment pathway.

GUIDELINES FOR FORMULARY EVALUATIONS

A commitment to NGS in therapy choice presents formulary committees with two decisions: (i) the choice of NGS platform for approval and reimbursement to support therapies for target populations within a disease area and (ii) the adoption of NGS sequencing to support the approval and reimbursement of new drug products to complement therapy choices for target populations with a disease area.

The premise that is fundamental to these *Guidelines* is that given any NGS platform an unrestricted remit, an open season ticket, for application in a health care system is not only to throw the concept of evidence based medicine out of the window, a scientific defenestration, but to open a Pandora's Box of potential clinical and budget impacts over which the health care system has no control.

Given this commitment to NGS as the 'gateway' to therapy choice, there are a number of unresolved issues which have to be addressed before approving an NGS platform and ensuring that submissions to support new drug products are presented in the context of NGS sequencing for target populations in specific disease areas. A targeting that is consistent with the approved indication for the product. These include:

- (i) choice of NGS platform
- (ii) barriers to integrating evidence to support NGS-based therapy choices in treatment guidelines
- (iii) implications of NGS for drug development and the assessment of single compound or single molecule claims
- (iv) impact of comorbid disease states
- (v) standards that formulary committees should adopt to evaluate NGS claims
- (vi) criteria for the pricing and approval of an NGS test
- (vii) the party responsible for making a submission to a formulary committee for an NGS platform approval.

The question of the ownership of an NGS platform is a critical issue in formulary submissions. Certainly, if current policies by the Centers for Medicare and Medicaid services (CMS) and the Food and Drug administration (FDA) are followed, the approval of a test or medical device stands alone²⁶. As long as the criteria for analytical and clinical validation are met, the question of robust evidence for clinical utility is put to one side. A physician can order a NGS test, receive the findings and then decide whether or not to act upon the findings on a case-by-case basis. But how is a formulary committee to justify the cost of the test? It can't be on standard clinical or cost-effectiveness criteria because the benefits of the test are specific to its performance in matching multiple mutation clusters to multiple drug combinations (some of which will be 'off-label' for that tumor indication) in the target patient population.

To what extent, for example, should an assessment of the clinical impact and cost-effectiveness of each of the individual therapy pathways prior to an 'overall' assessment of the platform performance be undertaken? Should each therapy pathway (or at least the

GUIDELINES FOR FORMULARY EVALUATIONS

mutation clusters with the expected greatest frequency) be assessed in the target population in terms of its respective claims being credible, evaluable and replicable? To provide a model scenario for claims in each of the disease areas and tumor stages where the NGS test is likely to be used is simply out of the question. In any event, with the continual updating of NGS platforms as our understanding of the genetics of a disease grows, together with the process of tumor growth and mutation diversification, trials would lack any long term credibility.

A formulary committee, therefore, would be well justified in rejecting a claim to reimburse an NGS platform because it has the potential to open a Pandora's box of mutation cluster driven drug choices which have, as noted above, not only an unknown clinical impact but also an unknown impact in terms of the cost of therapy for the target patient population.

A manufacturer, in this scenario, would have possibly little interest in sponsoring a specific platform to accompany a product submission with claims for the product driven by a basket trial matching therapy options to tumor mutation distributions. At the same time, a single compound claim for cost-effectiveness by the manufacturer for a new drug targeted at a single mutation or mutation cluster in the same patient population would be rejected on the grounds that it provides no guidance as to how that compound would fit into NGS driven treatment practice. In the absence of an evidence base that provides justification for the compound in the context of an NGS assessment, a new product submission would be for one compound among many that may be linked to specific mutations or mutation clusters. The independent contribution of that new compound to therapy specific and overall outcomes and costs would be impossible to assess. Claims made for the new compound would have no evidentiary base.

Given long standing concerns regarding the inability to replicate claims from phase 3 clinical trials, formulary committees are wise to be wary of claims based upon the application of NGS platform driven therapy choices. The evidence standard for NGS trials and observational studies (see section 1.5 below) should be no different from those applied to the 'classical' randomized controlled trial (RCT). This underscores the importance of a commitment by those making NGS-based submissions to post-approval studies. Indeed, given the number of potential target populations for an NGS platform, the need to demonstrate that the platform has clinical utility in the range of potential target populations is all the more pressing.

1.2.1 Next Generation Sequencing and Treatment Guidelines

If a case is to be made to a formulary committee for approving and reimbursing an NGS platform, with the emphasis on the need for a comparative assessment of competing NGS platforms, this it should be in terms of the potential place for the proposed NGS platform within the treatment guidelines for the target population. It is only within this framework that a coherent clinical and cost-effectiveness case for the platform can be made. Presenting a case only in terms of analytical and clinical validity is insufficient.

Ideally, an application for approving and reimbursing an NGS platform should follow the

GUIDELINES FOR FORMULARY EVALUATIONS

acceptance of that platform as integral to a treatment guideline. This will define a target population and the evidence base. The company making the submission should detail not only the place of the platform as part of the treatment pathway and stage of disease for a target population but also whether or not the platform has been approved by those developing the guidelines. Given the emphasis on evidence quality by organizations such as the National Comprehensive Cancer Network (NCCN), the submission should detail, for example, the evidence evaluated for the platform by the respective oncology panels, the assessed evidence quality and the most current recommendations for the platform as part of the therapy workup²⁷.

It is important to emphasize that it is difficult if not impossible to make a claim for the cost-effectiveness and justify the assessment price for a specific NGS platform in the absence of evidence for its impact in target populations. Unless evidence is presented, against competing platforms or a notional standard of care, that the application of the platform in target patient populations has generated improved outcomes and/or reduced costs there is no basis for pricing negotiations or even acceptance. In the absence of a robust evidence base detailing the impact of introducing the platform across a number of target populations defined by disease and stage of disease is presented, the health system runs the risk of introducing a device that increases costs with no evidence for improved outcomes. In the absence of a recommendation for a platform as part of a treatment guideline it is difficult to see whether claims for an NGS platform are to be seriously considered for reimbursement by a formulary committee.

1.2.2 NGS and Risk Stratification

NGS has the potential for risk stratification in target patient populations. Specific mutations or mutation clusters may be indicative of the risk of disease progression. In these cases there may be an option on whether or not to delay therapy choice until progression of the disease. Again, as a risk stratification tool, the application needs to be evidence based and linked to clinical assessments and other prognostic markers. If this is a feasible approach then there may be the possibility of introducing NGS platform evaluations at early stages in disease progression with the possibility of complete remission.

1.2.3 Barriers to NGS Implementation

At this juncture it is unclear as to how widely and how quickly NGS will be taken up. In the absence of a robust evidence base, for the clinical and cost-effective contribution of NGS the take-up is likely to be slow. Apart from the importance of distinguishing the various NGS platforms to identify those that meet (yet to be established) minimum standards for process and performance, the most promising avenue is probably through a focus of NGS platforms in late stage cancers. Situations where there are few options, a reliance on palliative care, and where irrespective of therapy options the patient has a poor survival and quality of life prognoses. Even so, the premium is still on the commitment by NGS platform manufacturers and, to a lesser extent, pharmaceutical manufacturers and biotechnology companies to commit to an investment

GUIDELINES FOR FORMULARY EVALUATIONS

in robust basket trials and observational studies. Unless a high quality evidence base is available, health guideline panels are unlikely to agree on the place of a specific NGS platform in therapy. Whether NGS manufacturers will commit to this is a moot point.

1.3 Acceptance of Modeled Claims

Modeled claims for cost-effectiveness, to include cost-minimization, cost-benefit and cost-utility have been the mainstay of consultant and academic health technology activities over the past 25 years. Putting to one side the question of whether, at least in the US modeled claims have had a significant impact on formulary decisions, the more fundamental question is whether modeled claims have a role in NGS submissions.

As noted above, one of the hallmarks of modeled claims is that they typically lack credibility. They are presented, at least for chronic disease states, in a lifetime simulation framework that purports to mimic the natural course of the disease. Yet at the same time the model fails to generate (by its design) any evaluable or replicable claims. While formulary committees in the US are perfectly at liberty to base formulary decisions and pricing on the constructed evidence of from models, on the construction of imaginary worlds, the uptake of NGS- based models presents challenges that argue against the continued acceptance of non-evaluable modeled claims.

This challenge is seen most clearly in claims for NGS platforms that argue for the cost-effectiveness of one platform compared to another. It is one thing to compare NGS platforms on their ability to identify mutations and mutation clusters; it is another to claim that the matching of mutations to therapy pathways generates more cost-effective outcomes compared to the therapy matching in another platform. At present we have no idea on the comparative performance of the competing NGS platforms or how future NGS platform entrants will perform.

A major obstacle to formulary acceptance of modeled claims driven by an NGS platform is that if the model is focused on a single compound, and even if targeted to a single molecule, the claim lacks credibility because of its limited application. The claim not only lacks external validity but fails to account for the potential wastage in allocating patients to a therapy where as many as 75% of those introduced may fail to respond; a failure potentially due to a neglect of the molecular basis of the disease in the target population. This does not mean that single compound models, whether they are molecule specific or merely targeted at the clinical manifestation of a disease, are necessarily put to one side. Rather, in basing a formulary submission on a single compound, those making the submission will have to justify why their approach is a satisfactory basis for formulary assessment.

1.4 Categorizing Formulary Submissions

One size does not fit all. With the advent of NGS platforms, formulary committees have to consider the evidence requirements and the questions they should address for the types of submission likely to be received. This applies both to submissions for new compounds or NGS platforms as well as to submissions requested as part on ongoing disease area and therapeutic reviews. Given the importance of maintaining NGS platforms to reflect accumulating evidence

GUIDELINES FOR FORMULARY EVALUATIONS

for the role of mutations in target populations and the need to track changing mutation expression as a tumor evolves, a regular review of approved NGS platforms is essential.

It should be emphasized that these *Guidelines* do not support, nor do they recommend, the owners of an NGS platform, seeking an ‘open’ approval for the application of an NGS platform across target populations in disease states. The application or approval of an NGS platform must be disease specific; it must be based on evidence specific to proposed target populations within disease states. Accepting the ‘specificity’ of competing NGS platforms provides the metric for single compound submissions. The submission, in proposing claims for the place of the compound in therapy, will need to demonstrate that the NGS platform used in the trial to match the compound to mutation clusters generates the same distribution of mutation clusters as the NGS platform already approved by the health system for the target populations in that disease state. Given this, the manufacturer can then proceed to evaluable claims that the new compound offers a clinical and/or cost benefit in those pathways where the compound is present either as a single therapy or in combination with other therapies.

Four categories of formulary submission can be considered:

1. **Standard Submission:** a submission which is focused on a specific compound targeting a clinically defined target population or disease state;
2. **Single Molecule Submission:** a submission which is targeted to a population defined by a specific molecular mutation (or group of associated mutations) within one or more disease states;
3. **NGS Platform Submission:** a submission to support reimbursement for a specific NGS platform in a target patient populations
4. **NGS Therapy Target Submission:** a submission that supports the application of an NGS platform in one or more target populations

For any of these four submissions to be accepted the standards for credibility, evaluation and replication of claims will apply. This implies that, unless indicated otherwise, all submissions must include a protocol to detail how the claims to support the submission are to be evaluated in the framework of a nominated NGS platform. This may involve a prospective observational study, a retrospective evaluation of a proposal for a phase 4 clinical trial. In each case, the focus must be on feedback to the formulary committee in a meaningful time frame. All claims must, of course, be comparative. This applies equally to the test or platforms utilized to identify specific mutations or mutation structures (e.g., the analytical and clinical validity of tests to support a targeted single molecule on NGS platform submission).

1.4.1 Standard Submissions

Standard submissions have, historically, been the mainstay of formulary submissions. The latest version of the AMCP Guidelines for Formulary Submissions and the NICE Guidelines exemplify this approach. The modeled claim is built around the expected or awarded indication for the product. The target population is defined in clinical terms, supported possibly by a diagnostic

GUIDELINES FOR FORMULARY EVALUATIONS

test, a checklist of symptoms or a series of patient reported (PRO) instruments intended to categorize the patient. The modeled claim for the compound extrapolates from RCT outcomes, supported by an indirect comparison if no head-to-head clinical outcomes are available, with the model extending over the lifetime of the patient cohort. Cost-outcomes claims are presented as constructed incremental cost outcomes ratios or probabilistic acceptability analysis with due account taken of structural and parameter uncertainty. As noted already, the claims presented in standard modeled claims typically fail the scientific standards of credibility, evaluation and replication. The model, given the protocol driving the RCT claims, takes no account of the genetic basis of the target disease. Although groups such as ICER are still wedded to this approach, the advent of NGS casts doubt on the validity both of the methodology and the information content of the claim.

If a submission is made for a single compound, irrespective of whether or not it claims to target a specific molecule or mutation, the case to be made should still be within the framework of a basket trial design where the compound is evaluated to target specific mutations. The days are long gone when a single compound case is made, with attendant lifetime cost-per-QALY claims, where the 'hit rate' is less than 50% (or even 30%).

1.4.2 Single Molecule Submissions

These lay claim to the potential for a higher 'hit rate' than the standard submission model, but suffer from the same weaknesses in the limited scope of the evaluation and the neglect of more complex mutation loads and the distribution of mutation clusters across the target population. Indeed, they limit the scope of the compound in reporting only that the presentation of the target compound is likely to yield positive results. No account is taken of the possibility of associated mutations qualifying claims. Again, it is possible to model from the base case RCTs for the intervention compound, although the modeled claims would typically lack credibility, evaluation and replication. A further concern is with the accuracy of the test in identifying a target population group.

Once again, if the compound is focused on a single molecule, the case for the product still has to be made (as in the case of a standard submission as detailed in Section 1.4.1) within the framework of a mutation distribution driven multi-therapy pathway model. The formulary committee has to be convinced that the new compound adds value in terms of outcomes and costs as an addition to those compounds presently identified for the various therapy pathways.

1.4.3 NGS Platform Submissions

The key issue for those making an NGS platform submission is the need to establish the platform's merits against those of competing NGS platforms within individual disease areas. Certainly, a claim can be made that an NGS platform meets the required standards for analytical and clinical validity; the question is whether the platform has clinical utility in target patient populations. That is, it provides an acceptable framework for identifying mutation load and mutation clusters in that population.

As noted above, two critical issues are (i) the place of the NGS platform in treatment pathways and (ii) whether the platform has been accepted as an integral part of treatment guidelines.

GUIDELINES FOR FORMULARY EVALUATIONS

However, even if an NGS platform has been accepted as the ‘gold standard’ platform for that disease-specific pathway, the submission would still need to demonstrate that the evidence base for the acceptance of the platform has been replicated for those target populations. As well, this would not relieve those making the submission for the platform from developing and underwriting a protocol to establish the claims for that platform in the target population. After all, the evidence base for all too many treatment pathways is all too often considered to be barely adequate to support therapy decisions.

In making a submission, the committee would expect the clinical utility claims to be supported by reference to the literature for the designated target population(s). At the same time, the submission should include a protocol to establish the value of the platform in the target population(s). Again, this should be in comparative terms and should include any NGS platforms already approved for that target population.

1.4.4 NGS Therapy Target Submissions

Ideally, a submission to a formulary, where there is an established molecular basis for a disease, should combine an NGS platform with comparative modeled claims for the impact of therapy choice on the target population. As long as the modeled claim is designed to generate feedback to a formulary committee, those making the submission are at liberty to develop the model they think most appropriate for the target patient population. Ideally, the comparison should be with (i) the notional standard of care and (ii) a comparator NGS platform. Again, this should be in comparative terms and should include any NGS platforms already approved for that target population.

1.5 Quality Adjusted Life Years

Under these *Guidelines*, quality-adjusted-life year (QALY) claims are only acceptable if (i) they are expressed in evaluable terms; (ii) a protocol is presented detailing how the claim is to be evaluated and reported to a formulary committee in a meaningful timeframe and (iii) the submission justifies the use of the particular QALY measure for the disease area. Modeled or simulated claims that track the course of a chronic disease, for example, are of little interest as there is no chance that they could be evaluated. They have no credibility. A similar objection applies to claims expressed in lifetime cost-per-QALY terms that apply notional cost-per-QALY thresholds to justify premium prices or discounting of competitor pricing. Although a QALY endpoint has been proposed as the ‘gold standard’ outcome claim, the fact is that in the US QALYs are never collected as part of health care systems’ administration. Formulary committees have expressed little interest in QALY based claims as a useful input to decision making and even less interest in willingness-to-pay thresholds.

1.6 Submission Protocol

The purpose of these *Guidelines* is to ensure that claims supporting formulary decisions are credible, evaluable and replicable. Manufacturers and others who submit claims are asked to justify the claims made by presenting those claims in an evaluable form, supported by an assessment protocol. The assessment protocol is central to the assessment of claims. It requires manufacturers to stand behind their product and ensure that the evidence base supports claims

GUIDELINES FOR FORMULARY EVALUATIONS

replication. Until an assessment has been made, claims made for products should be treated as provisional. The decision then rests with the formulary committee either to accept provisionally a protocol subject to the claims being evaluated and reported back to the committee or to put a formulary decision to one side until a claims assessment has been presented. If an assessment has already been made then these results can be reported to the committee as part of the NGS submission.

In these guidelines, while there is no hard and fast format for making an NGS submission, there are questions that a formulary committee should ask when a submission is received. Questions are proposed for submissions where claims are from (i) standard or single molecule submissions; (ii) a submission to support a single NGS platform; and (iii) a submission to support the impact on costs and outcomes of therapy pathways driven by an NGS platform.

1.7 Target Populations

A key element in the claims protocol is the identification of target populations. With the increasing emphasis on personalized or precision medicine, the expectation is that diseases will be defined in terms of their molecular taxonomy or pathology rather than their symptoms, gross pathology or part of the body in which they appear²⁸. At the same time, those arguing for a target population need to justify the choice of an NGS platform as a decision support tool. The choice of a diagnostic test or device to identify the target population must meet regulatory standards. Replication of claims, which will likely become more necessary as manufacturers compete for smaller target populations with high-cost products in the same molecular ‘space’, will have to be able to identify target populations for such assessment. At the same time, asking target populations to be identified in a protocol will contribute to defining and establishing the boundaries of a disease state for claims evaluation and replication. This, hopefully, will limit the extent to which ‘disease mongering’ occurs²⁹.

1.7.1 Drug and Platform Costs in Target Populations

With increasing concern being expressed over the pricing of drugs, particularly in the anticancer area for end-of-life metastatic treatments, a critical input to a formulary decision is the launch price (wholesale acquisition cost) of a product and the current price of comparator therapies for the target indications³⁰. These *Guidelines* ask manufacturers to indicate what the anticipated launch price of the drug is expected to be (or the current price if it has already been launched). Where a drug has been launched, information should be provided on whether or not there have been price changes since launch. At the same time, for comparison purposes, manufacturers are asked to provide anticipated launch or current prices in other global markets: Canada, the European Union and Australia. The same requirements apply to NGS platforms. What is the ‘list’ price? Does the list price vary by mutation scope? How has the list price been justified? If the NGS is accepted in other markets, what price has been negotiated?

As well as launch price, manufacturers are asked to provide estimates of the costs of an episode of care for their product compared to comparator products. This can be expressed as the cost over an expected survival period for end-of-life therapy or the annual cost in an ongoing chronic disease intervention. The episode of care costs should capture (i) drug costs and (ii) total direct medical costs for patient evaluation, monitoring and administration. Where an NGS platform supports linking mutations to therapy pathways, then the submission should incorporate (i)

GUIDELINES FOR FORMULARY EVALUATIONS

therapy pathway resource utilization and costs and (ii) outcomes for each therapy pathway. Once these data are to hand, then a submission should include overall costs for the target population and the outcomes profile. The submission should also indicate (i) where there are no therapies matched to a mutation or mutation cluster; (ii) incomplete matching of mutation clusters to therapies; and (iii) therapies that are considered ‘duds’ in the matching of mutations to specific drugs.

When modeled claims are presented care should be taken to include the cost and frequency of patient assessment through tests or other assays. In late stage cancer, the potentially rapid progression of the disease may require ongoing reassessments utilizing the NGS platform and therapy modifications. The modeled claim should include criteria for reassessment and the ongoing assessment costs.

1.8 Options in Claims Evaluation

The credibility of claims to support formulary assessments of pharmaceutical products and devices can be established by:

- A retrospective assessment of established products as inputs to ongoing disease area and therapeutic class reviews
- A protocol-driven prospective or tracking assessment of prospective claims for new products or devices matched to existing products
- A protocol-driven prospective observational assessment of claims in the context of a phase 4 trial or similar

The standards proposed here apply equally to all modeled submissions. While it would be possible to undertake retrospective studies for standard modeled claims, the likelihood is that any claim presented will have to look to a prospective study design to evaluate the credibility of claims. This assessment would be driven by a protocol which should accompany a claims submission. In the case of NGS driven therapy target submissions, the claims should be presented in terms of both the individual pathways defined by the matched therapy and mutation distribution (e.g., pathway median survival, direct medical costs) as well as the overall impact of the NGS platform on the ‘aggregate’ outcomes achieved for the target population (e.g., overall median survival, overall direct medical costs). The latter measure provides the basis for comparison with the standard of care and comparator NGS platforms in the target population (e.g., stage 4 metastatic melanoma patients).

1.8.1 Comparative Effectiveness Research

The standards for formula submission recommended here are consistent with interventions under, for example the *Patient Protection and Affordable Care Act* to support quality outcomes and metrics in health care delivery. The standards are designed to provide a robust evidence base for formulary decisions in target patient populations. These populations may be risk stratified to match the proposed indication for the pharmaceutical product. The standards proposed can support targeted interventions to improve clinical outcomes in, for example, at-risk populations.

Implementation of these *Guidelines* proposed protocols to support the assessment of comparative clinical and cost-effectiveness claims is, of course, consistent with the objectives of comparative

GUIDELINES FOR FORMULARY EVALUATIONS

effectiveness research. The principal difference is that these Guidelines link the assessment to claims made for the product on market entry and not an assessment sometime after market entry. This may involve, not a prospective observational study to track product uptake and response, but also a randomized controlled trial. Once the product is listed on formulary there may be less incentive or willingness to evaluate comparative clinical and cost-effectiveness claims. Linking a claims assessment protocol to a product submission for formulary listing, with a commitment to underwrite a comparative evaluation, provides such an incentive.

1.9 Value Claims

Putting on one side claims couched in QALY terms also means that, unless specifically requested by a formulary committee, attempts to construct composite cost-outcomes claims are also unlikely to convince formulary committees. While it might seem paradoxical after some 20 or more years focusing on composite outcome measures such as QALYs that allow, at least in the context of reference case or similar models, gross comparisons across disease areas, the view is now that value claims are best seen as stakeholder specific and within disease areas.

The absence of agreements on disease specific outcome metrics and on metrics that might allow comparisons across disease areas means that it is the responsibility of health systems to establish standards for evaluating comparative drug and NGS platform claims. These might involve health system and disease area specific value frameworks such as those proposed by the American Society of Clinical Oncology or estimates of the impact of a new therapy on resource utilization and health system budgets³¹. Irrespective of the value framework, the two key elements are that: (i) a value framework should encompass a range of outcome measures (which a health system may weight to create a composite index) and (ii) the value measures should be credible and evaluable.

The choice of a value framework is at the discretion of stakeholders in the health system. These *Guidelines* are not concerned to set value standards but rather to ensure that stakeholders seen that value claims are provisional and necessarily subject to evaluation and replication. It is entirely at the discretion of the health system to indicate which outcome measures are appropriate for claims submissions and assessment protocols. These could include quality of life, patient satisfaction with therapy choice, toxicity or even willingness to pay. The only criterion is that the claims are credible, evaluable and replicable.

1.10 Adherence and Persistence

Claims from models and simulations typically make no allowance for patient adherence and persistence following product launch. This is an unfortunate oversight as medication adherence and persistence is seen as a global health problem. Despite decades of research the patterns of poor adherence and lack of persistence remain essentially unchanged. This is true across all major disease areas and most notably for older patients with chronic disease.

Given the potential impact of poor adherence and the lack of persistence on clinical, cost-effectiveness and other value claims, the GFE proposes that where a submission is prepared it should not only provide a systematic review of adherence and persistence behavior for the target population and comparator therapies within the disease or therapy area but should also detail, if considered relevant, potential interventions to mitigate such behavior. This is particularly

GUIDELINES FOR FORMULARY EVALUATIONS

relevant where therapy choices are driven by an NGS platform. The presence of mutation clusters for individual patients may lead to recommendations for combination therapy (which is likely to include ‘off-label’ compounds for the target tumor). The likelihood that combination therapy pathways, together with therapy for comorbidities may overwhelm the older patient is a real concern. Unfortunately, there is a dearth of data tracking adherence and persistence for patients in these groups (e.g., likelihood of adverse events triggering discontinuation). Indeed, claims for the impact of an NGS platform driven therapy choices should include claims for adherence and persistence behavior. This should be addressed for the individual patient pathways as well as for the overall impact of in the target population.

1.10.1 Physician Uptake

Claims for the overall benefits of an NGS platform linked therapy package in target patient groups rely, not only on patient adherence and persistence with therapy, but also the extent to which treating physicians order and assay and, with the results to hand, implement the recommendations linking tumor mutations to therapies. Again, there is a dearth of evidence to capture either of these actions. Once again, submissions should be explicit as to the claims for NGS platform penetration in the target physician population: (i) the proportion of physicians expected to order an assay; (ii) the proportion of these who are expected to implement the NGS platform recommendations; and (iii) expected physician adherence and persistence with the therapy. The claims are readily evaluable.

In addition, the submission should include a description (with examples) of the material sent to the physician describing the results of the assay, its implications and outcomes for the individual patient and the implementation of the targeted therapies. Ideally, the submission should be able to report on how the messaging was developed and the benefits, if any, of varying message content by specific therapy combinations (e.g., likelihood of adverse events and discontinuation by the patient).

1.11 Comorbidities

Older patient population typically present with one or more comorbidities (e.g., cancer with diabetes, cardiovascular conditions). From a therapy perspective where choices are matched to mutations, a question that a formulary committee needs to address is the scope of the NGS platform assessment? Is the assessment to focus on a specific tumor type and put to one side the presence of other disease states? If these disease states also reflect germline or somatic mutations, the formulary committee may request a complete sequencing to capture targeted therapy options in these disease states. If so, this request should be part of the proposed claims assessment protocol.

Accommodating comorbidities in patient management adds a further level of complexity. Therapy choices for comorbidities may not only lead to potential drug-to-drug interactions, but the success or otherwise of a therapy choice linked to a mutation cluster in the target disease state may adversely impact disease progression in a comorbid disease state. The answer may be to extend the modeled claim to include principal comorbidities and the options for targeted therapy choices for those comorbid states.

1.12 The Evidence Base

Irrespective of the type of submission, those submitting must meet established standards for systematic literature reviews and the evaluation of published studies. The key standards are:

1.12.1 Diagnostic Accuracy and Quality

All studies referenced to support claims for diagnostic accuracy for the target test and comparators should be evaluated against the STARD 2015 (Standards for Reporting of Diagnostic Accuracy Studies) statement³². Each study reported should be scored. Also, all studies referenced to support claims for diagnostic accuracy should also be appraised for quality and bias against the QUADAS-2 tool (Quality Assessment of Diagnostic Accuracy Studies)³³. Each study reported should be scored.

1.12.2 Systematic Reviews

Whenever a systematic review is requested to support the test and comparator evidence base it must conform to the PISMA-P 2015 standards [<http://www.prisma-statement.org/Extensions/Protocols.aspx>]³⁴. Apart from the usual databases (e.g. PubMed), reference should also be made for reviews focused on the CDC Public Health Genomics Knowledge Base³⁵.

1.12.3 Reporting Randomized Trials

Reporting of results from randomized clinical trials of test performance should conform to the Consolidated Standards of Reporting Trials (CONSORT)³⁶ [<http://www.consort-statement.org/>]. This is a standard format for reporting on trial organization, analysis and interpretation. The CONSORT Statement comprises a 25-item check list [<http://www.consort-statement.org/checklists/view/32-consort/66-title>] and flow diagram [<http://www.consort-statement.org/consort-statement/flow-diagram>] to record the progress of patients through the trial.

1.12.4 Evidence Hierarchy

Claims for the efficacy or effectiveness of tests in clinical practice must be founded on high quality and bias-free evidence. Where a submission has undertaken a systematic review or relies upon individual studies to support credible, evaluable and replicable claims the evidence presented should be assessed against the standards established within the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working groups. The GRADE framework has superseded earlier proposals for the ranking of evidence (which typically ranks from randomized trials through to observational studies and anecdotal, key opinion leader evidence) to a more flexible evidence hierarchy addressing the quality of evidence for individual outcomes. Specifically: bias, inconsistency, indirectness, imprecision and publication bias³⁷.

The GRADE framework is intended to apply to meta-analyses from systematic reviews but can be applied to individual studies or non-quantitative syntheses. The essence of the GRADE approach is that, within each hierarchy level, it allows the downgrading or upgrading of

GUIDELINES FOR FORMULARY EVALUATIONS

evidence. Downgrading, for example in the case of randomized clinical trials, occurs if there is a risk of bias, inconsistency, indirectness, imprecision and publication bias. Upgrading, for example in the case of non-randomized studies can occur if there is a large magnitude of effect, evidence of a dose response effect and if all plausible confounding factors have been taken into account. The application of the GRADE framework 4-level quality rating hierarchy is detailed in the Cochrane Collaboration handbook³⁸. The ratings are:

1. *High Quality Rating*: Randomized trials; or double-upgraded observational studies
2. *Moderate Quality Rating*: Randomized trials; or upgraded observational studies
3. *Low quality rating*: Double-downgraded randomized trials; or observational studies
4. *Very low quality rating*: Triple-downgraded randomized trials; or downgraded observational studies; or case series/case reports.

The GRADE evidence approach has figured largely in the Agency for Healthcare Research and Quality *Methods Guide for Comparative Effectiveness Research* to support the Evidence-based Practice Center (EPC) Program³⁹. The EPC framework grades the strength of evidence from RCTs as well as observational studies in a systematic review through assessing specific domains: study limitations, directness, consistency, precision and reporting bias. Potential additional domains are: dose-response association, plausible confounding for observed effect and strength of association. Scoring these domains yields four strength of evidence grade:

1. *High*: The reviewers are very confident that the estimate of effect lies close to the true effect
2. *Moderate*: The reviewers are moderately confident that the estimate of effect lies close to the true effect
3. *Low*: The reviewers have limited confidence that the estimate of effect lies close to the true effect
4. *Insufficient*: The reviewers have no evidence, they are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome

The importance of evidence standards in precision medicine claims should not be understated. If the application of NGS (and presumably the generations after that) is to be accepted it is crucial that we have publicly accessible, comprehensive, structured, transparent and bias-free evidence that support both claims development from analytical and clinical validations but subsequent clinical utility claims; claims that are credible, evaluable and replicable. Comparative and systematic reviews are central to this process: comparative because the choice of a test and a belief in the validity of the test is critical to its acceptance by treating physicians as an integral part of the process of care, and systematic because we require the accumulation of evidence from well conducted and credible effectiveness studies to support patient involvement in treatment decisions.

1.12.5 Evidence Base in Oncology

GUIDELINES FOR FORMULARY EVALUATIONS

Given the potential impact of NGS in oncology, it is worth noting the NCCN categories for evidence and the degree of consensus among guideline panel members evaluating the evidence base for therapy decisions. These are:

Category 1: Based upon high-level evidence, there is **uniform** NCCN consensus that the intervention is appropriate;

Category 2A: Based upon lower-level evidence, there is **uniform** NCCN consensus that the intervention is appropriate;

Category 2B: Based upon lower-level evidence, there is NCCN consensus that the intervention is appropriate;

Category 3: Based upon any level of evidence, there is major NCCN disagreement that the intervention is appropriate.

All recommendations in the NCCN guidelines are category 2A unless otherwise noted in the treatment guideline. All too often, unfortunately, the evidence is category 2B.

GUIDELINES FOR FORMULARY EVALUATIONS

2 EVALUATING FORMULARY SUBMISSIONS: KEY ELEMENTS

2.1 Product and Comparator Descriptions and Pricing

Typically, where a submission is made to a formulary committee for a single product or individual compound, either under a standard submission or molecular targeted submission for a single compound, the manufacturer would be expected to provide detail on the indication sought and the package insert for the product and its comparators together with any black box warnings. The choice of comparators needs to be justified. Where approval is being sought for an NGS platform the submission is more complex. First, there should be a description of the platform; second, a statement of the target disease states for which approval is being sought; and, third, comparator NGS platforms. Approvals for the platform (e.g. as a medical device) should be provided.

In the case of individual compounds, the manufacturer should submit the Wholesale Acquisition Cost (WAC) and a statement as to the company policy price increases over the patent life of the product. Similarly, if an NGS-platform is involved, details on the price (or price list) for the platform should be provided with, again, company policies regarding price increases. Corresponding prices for comparator compounds or NGS-platforms should be provided.

In the case of an NGS platform targeted to a specific disease start, the submission should include a list of the compounds most frequently identified as components of mutation driven therapy pathways, their current indications and the WAC for each compound.

2.2 Target Population and Place in Therapy

For each individual compound or NGS platform the submission should include a description of the target patient population(s), including target sub-populations, their molecular profile for the target disease and the place of the product(s) in therapy. If there are agreed national treatment guidelines these should be provided indicating where there may be anticipated changes following approval of the new product. The descriptions of the target populations should include a profile of co-morbidities and the prevalence of those co-morbidities, together with their molecular profile and the most frequently identified compound combinations in mutation driven therapy pathways..

It is important to remember that all claims for interventions in a target population have to be seen in the context of a NGS platform for that target population. The NGS platform is critical as it has the potential to identify the distribution of mutations as therapy targets within that population.

2.3 Significant Adverse Events and Contraindications

For each of the individual compounds and comparators a summary should be provided detailing the significant adverse events ($\geq 1\%$) and contraindications. The adverse event profiles should detail (i) those reported as part of the phase 3 trials and (ii) those reported as part of risk assessments following product approvals.

GUIDELINES FOR FORMULARY EVALUATIONS

In the case of NGS platforms supporting therapy intervention pathways, the submission should include, by therapy pathway, the most common adverse events associated with those pathways.

2.4 Primary and Secondary Outcomes

For each of individual compounds and comparators a summary should be provided detailing for the respective pivotal phase 3 trials the primary and secondary outcomes. This summary should detail if the claims made for the product and comparators are consistent with the primary outcomes identified for the phase 3 trials for the therapy pathways within the target population. Each pathway should be defined in terms of a common primary endpoint, together with clinically significant secondary endpoints and those defined by patient reported outcomes.

If, during the trial process, the primary outcomes were switched and/or secondary outcomes dropped/added this should be documented. A statement is required indicating that this has not occurred in any of the product and comparator trials. If outcome switching has occurred a statement is also required that the trial was powered to support those claims. Summaries of the pivotal trial protocols should be provided as part of the submission.

In the case of NGS platforms, the submission should indicate the common primary outcome across all pathways (e.g., median survival time) together with the appropriate secondary outcomes. These may include clinical markers as well as patient reported outcomes.

2.5 Direct and Indirect Comparisons

If there have been head-to-head comparisons between individual compounds and comparators these should be documented. Similarly, the results of indirect comparisons should also be documented, together with a brief description of the techniques employed. If the claims from indirect comparisons have been evaluated from phase 4 trials or observational studies these should be documented.

Once again, as the focus of the submission is on therapy pathways defined by the NGS platform, head-to-head comparisons between competing platforms in target patient populations should be presented. If the competing NGS platforms have reported outcomes in similar target populations the equivalent of an indirect comparison between platforms including mutation distributions, therapy pathway specifications and outcomes by individual pathway and overall for the platform across all pathways should be presented.

2.6 Replication

For each of the individual compounds and comparators attempts to replicate the primary outcomes of the phase 3 trials in successor trials or observational studies these should be documented together with the respective results. If there is a lack of concordance between the original phase 3 outcomes and those reported for any replication this should be reported.

Replication is also at the center of the justification for adopting an NGS platform. In this case, with the number of mutation complex driven treatment pathways for a target treatment population, formulary committees need to set standards for NGS platform assessment together with the outcomes for patients allocated to specific treatment pathways. There needs to be an ongoing process of claims review, driven in part by the molecular changes in tumors as a disease

GUIDELINES FOR FORMULARY EVALUATIONS

progresses but also by the ability of NGS platforms to update their libraries and potentially modify their assessment of mutation load and the distribution of mutation complexes and the choice of target drugs.

For the NGS platform being proposed and the comparator platforms, the results of replication studies across the target patient populations should be reported.

2.7 Adherence and Persistence

For each of the product and comparators evidence is required for adherence and persistence behavior (i) during the pivotal phase 3 trials and (ii) for a period up to 4 years from product launch. This review of adherence and persistence should report on descriptive studies as well as those that have attempted to assess the determinants of this behavior in the therapeutic product class. The review should report on the extent to which patterns of adherence and persistence have modified primary clinical claims for the product and comparators in the target population or sub-populations, together with the results of any interventions to impact adherence and persistence behavior.

Claims for and reported patterns of adherence and persistence specific to the therapy pathways defined by the NGS platform and its comparators with the target patient population should also be presented.

2.8 Completed Evaluations

If a manufacturer has already completed an evaluation for the clinical and cost-effectiveness claims it proposes to submit for the product or device then this should be included as part of the submission. The manufacturer should provide copies of the protocol and NGS platform that supported this evaluation. The manufacturer should state why this previous evaluation is appropriate for claims made for the target health system population and why replication of this previous study is unnecessary.

2.9 Evidence Summaries

The submission should include, as an Appendix, systematic review and evidence summaries in spreadsheet form for all individual compounds and the comparator randomized clinical trials, together with trials for the NGS platform and its comparators. The focus should be on basket trial designs where either individual compounds or therapy combinations are assessed. These summaries should conform to the standards detailed in Section 1.11

GUIDELINES FOR FORMULARY EVALUATIONS

3. VALUE AND COST-OUTCOMES CLAIMS

3.1 Scope and Options in Value and Modeled Cost-Outcome Claims

Value and cost-outcomes claims submitted are only considered if:

- The claims submitted are evaluable
- The proposed claims evaluation provides feedback in a meaningful time horizon

There is no restriction on the type of value or cost-outcomes claims that can be submitted. Value claims can be expressed in clinical terms, as patient reported outcomes, as direct medical costs and as cost-outcomes. Where patient reported outcomes are used they should meet accepted standards for their measurement properties with a statement indicating whether or not interval differences are clinically meaningful.

Claims presented for either individual compounds or NGS platforms should be evaluable within a 2-year time frame. Longer time frames, including those that model the natural course of a chronic disease are not relevant to formulary decision making. The objective should be to present claims that can be evaluated and reported back to decision makers to support effective decision making.

NGS platform claims assessment should be specific to individual therapy pathways for specific target populations. Claims for platform or product impact which fail to take account of the heterogeneity of tumor expression and the need to match therapies to mutation distributions should be rejected. Overall claims can then be generated by a weighted aggregation over the individual pathway claims

3.2 Claims Evaluation Protocol

The formulary submission should be accompanied by a protocol as detailed in Sections 4 and 5. From an evaluation perspective, it is important that claims made are capable of being assessed either from existing data sets (e.g., administrative claims linked to EMRs) to capture resource utilization impacts as well as clinical endpoints or from a prospective observational study to capture outcome that are not typically captured in EMRs. Those making the submission have the choice, therefore, of tailoring their claims to available data from third party vendors or underwriting a prospective study.

The claims evaluation protocol must include a preferred NGS platform to determine the allocation of patients by treatment pathway. The platform could be one that the health system has already determined is appropriate for that target population to support the formulary approval of a new compound or it could be a composite protocol that includes claims for a new platform and the impact of the indicated therapy pathways utilizing currently approved drugs.

If an NGS platform is being assessed then the protocol should focus on the claims for that platform versus competing platforms. The claims, as noted above, should be (i) specific to each therapy pathway over a defined time horizon and (ii) for the platform overall.

3.3 Value Claims and Model Structures

It is entirely at the discretion of those making the submission whether the value and cost-outcomes claims are presented as specific clinical endpoints, as outcomes from a simple decision model or as a more complex structure such as a Markov model or a discrete event simulation. It is entirely at the discretion of those submitting claims whether or not to present them as incremental cost outcomes ratios supported by sensitivity analyses. The model may attempt to capture the impact of an individual compound against comparators or, more usefully, it could attempt to generate evaluable claims for the therapy pathways matched to the most frequently observed mutations from an NGS platform. If the latter model is presented then it should be agreed with the formulary committee that the choice of NGS platform is appropriate for the target population and any compounds of particular interest, such as a new compound developed by a manufacturer, supporting evaluable claims in the context of the model.

Again, it is also entirely at the discretion of those making the submission whether they want to support their choice of model and the assumptions driving the model by reference to quality checklists such as CHEERS⁴⁰. The focus on generating testable predictions means that a model is not judged on the correspondence of its input assumptions and core mechanism to some perception of reality. The issue of the sufficiency of correspondence necessarily entailing the claims made is irrelevant. The merits of a model rest upon its ability to generate testable claims and the results of an assessment of those claims.

3.4 QALYs and Cost-Outcomes Thresholds

The *Guidelines* advise against claims being presented as utilities or as costs-per-QALY (quality adjusted life years) unless they can be generated and reported in a 2-year time frame. Apart from the obvious case of lifetime or long-term cost-per-QALY claims which are clearly unevaluable and which should be rejected as reflecting simulated imaginary worlds, the *Guidelines* would also caution that few health care decision makers in the US, in either public or private domains, have the slightest interest in QALY based claims, whether they are drawn from RCTs or are modeled at second hand from the literature. In the first place, there is no agreed standard for measuring QALYs, in the second place QALY measures are seldom (if ever) reported as the primary outcome measure in RCTs, and thirdly, they are never captured in electronic medical records or other data bases. In addition, it is not the intention of the *Guidelines* to support modeled cost-utility claims that are designed to demonstrate that a product or device is ‘cost-effective’ because it generates incremental claims that are below a notional community willingness to pay threshold. If a formulary committee is prepared to base its decision making on willingness to pay thresholds that is their decision and they should inform those making submissions accordingly.

If QALYs are considered an important endpoint for the target population, then they need to be captured across the various treatment pathways identified by the NGS platform. Given the range of QALY instruments, it is advisable that agreement is sought from a formulary committee regarding the choice of QALY instrument. The same requirement holds for claims assessment protocols that capture other patient reported outcomes.

GUIDELINES FOR FORMULARY EVALUATIONS

3.5 Adherence and Persistence in Modeled Claims

Modeled claims submitted should accommodate anticipated adherence and persistence behavior. In many chronic disease states product utilization data point to relatively low rates of adherence and persistence by the end of the first 12 months from the index prescription. Those presenting a submission should make clear what the anticipated adherence and persistence behavior for the product or therapy pathway is expected to be and whether or not it offers advantages in these respects over comparators. Claims should not be based on a possible minority of patients introduced to a new therapy who are assumed to be adherent (e.g., medical possession ratio \geq 0.8) and persistent over a modeled time horizon of, say, 2 years.

It is, of course, entirely at the discretion of those making the submission to suggest, as part of their claims for the product to propose evaluable intervention strategies that could support longer term adherence and persistence.

3.6 Dispensing with a Modeled Claim

It is not mandatory for those subscribing to the *Guidelines* for a modeled claim to be presented. If those preparing a submission believe they are not in a position to develop a credible, composite cost-outcomes modeled claim then this should be put to one side. The focus of the submission would be, therefore, on evaluable comparative clinical or value claims such as comparative cardiovascular events avoided or median survival as well as resource utilization, direct and indirect medical costs for patients and caregivers.

3.7 Concurrent Submissions

In situations where a manufacturer has already submitted a protocol to another health system and is underwriting an evaluation this should be reported as part of the submission. In this case, as long as the formulary committee agrees that the protocol is acceptable, then the manufacturer should commit to reporting to the committee once the other study is complete. There is no requirement for the evaluation to be undertaken in the health system's patient population.

GUIDELINES FOR FORMULARY EVALUATIONS

4 PRODUCT CLAIMS ASSESSMENT PROTOCOL

4.1 Protocol Standards (PROST)

The claims assessment protocol can be applied to clinical claims, modeled cost-outcome and budget impact claims for a specific product. The assessment checklist recommended by the *Guidelines* is a revised version of an earlier Protocol Standards (PROST) checklist⁴¹. The protocol is designed to support prospective studies where patients are tracked from existing data sources as well as prospective observational studies involving reporting from human subjects. The protocol can also support prospective applications of value calculators where, rather than relying on inputs from randomized clinical trials, the elements of the value calculator based on those trial results can be evaluated with potential feedback to support re-calibration of initial value scores.

This protocol is intended to apply single product applications for formulary listing, to include those products limited to targeting a specific molecule or mutation. As noted above, the assessment of claims made has to be in the context of an NGS platform where the new compound is assessed through its contribution to outcomes and costs within treatment pathways identified for target populations.

The protocol should identify the target group for the claims assessment. If the product is intended to apply to more than one target group (e.g., an NGS platform driving therapy choice across tumor types) then a separate protocol should be submitted for each target population. The target population matched protocol requirement would apply also to targeted sub-populations with a disease state where the NGS platform or product intervention may be recommended at different disease stages. If an NGS platform is the basis for target patient identification, the submission should also include the questions proposed in Section 5.

4.2 Format for Product Protocol Submission

Title: *Proposed Assessment Protocol for Evaluating Clinical and Cost-Outcome Claims for [product]*

Abstract: *Structured summary of protocol assessment objectives: (i) claims to be evaluated; (ii) target population; (iii) proposed study design; (iv) data sources; (v) timelines for evaluation; (vi) reporting format; and (vii) choice of NGS platform (if appropriate)*

Context: *A statement of the context of the study to include: (i) a systematic review of clinical and cost-effectiveness claims for the product and comparators in the relevant time horizon and, where appropriate, from application of NGS platforms in the target population; (ii) a systematic review of adherence and persistence patterns for the product and comparators in the target population; and (iii) a statement of the contribution this study is expected to make in evaluating competing therapy pathway interventions in the target population.*

NGS Platform Application: *If the protocol is concerned only with a single compound or a compound targeted to a single mutation or mutation cluster, the submission should detail how*

GUIDELINES FOR FORMULARY EVALUATIONS

product claims are to be assessed within an NGS platform driven specification of treatment pathways defined by the distribution of mutations.

NGS Platform Choice: *If the protocol includes the application of an NGS platform to target patients for therapy options, then there needs to be a statement that supports the choice of the NGS platform as opposed to potentially competing platforms.*

Claims: *A statement of the claims that the proposed study will be evaluating: (i) clinical outcomes; (ii) adverse event and safety outcomes; (iii) resource utilization outcomes; (iv) direct medical costs; and (v) cost-effectiveness. These claims should be specific to therapy intervention pathways.*

Target Population: *Characteristics of target population (age, gender, ethnicity, race) to include sub-groups with clinical and genetic markers (presence/exclusion of co-morbidities; stage of disease) and distribution of mutation clusters.*

Data: *Description of proposed data source(s) and any permissions required to access data/target population (e.g., IRB approvals, EMR approvals, access to administrative claims data) and confidentiality requirements*

Study design: *Rationale and description of study design, to include: (i) rationale for choice of study design in target population (e.g., choice of basket protocol); (ii) concurrence with good practice guidelines; (iii) treatment comparators; (iv) timeframe; (v) statistical and analysis plan; (vi) data sources; (vii) statistical hypotheses ; (viii) description of statistical methods; (ix) sample size; (x) procedures to minimize bias; (xi) quality assurance; (xii) confidentiality; (xiii) data handling; and (xiv) reporting.*

Budget: *A detailed budget and a commitment to underwriting the evaluation.*

Interim Reporting: *Proposed schedule for interim reporting of study implementation and progress.*

Final Report: *Proposed submission data for final report and proposed report outline.*

4.3 Questions a Formulary Committee Should Ask

On receipt of a protocol the formulary committee should consider:

Objectives: *Has the protocol provided a summary of the study objectives? Does the protocol abstract provide a summary of (i) population targeted; (ii) choice of NGS platform; (iii) comparators; (iv) timeframe; (v) study design; (vi) data sources; (vii) endpoints; and (ix) reporting format.*

GUIDELINES FOR FORMULARY EVALUATIONS

Context: *Has the protocol provided a systematic review to establish the context for the claims to be evaluated and the contribution this evaluation have to an overall assessment of comparative treatment benefits and harms in the population targeted and indication for the product?*

Target Population: *Has the protocol provided a rationale for the population to be targeted, to include (i) demographic characteristics; (ii) clinical status; (iii) disease stage; (iv) co-morbidities; and (v) prospective therapy pathways*

Claims: *Has the protocol provided a summary of the modeled claims for cost-effectiveness to be evaluated? Has the protocol identified claims for (i) clinical benefits; (ii) adverse events and safety; (iii) resource utilization; (iv) direct medical costs; (v) cost-effectiveness and (vi) adherence and persistence? Are the claims both specific to the therapy pathways and overall for the NGS platform and, where appropriate, place of the new compound in therapy?*

Study Design: *Has the protocol provided a rationale for the study design (e.g., prospective effectiveness trial, basket observational design, cohort/ observational study)? Were alternative study designs proposed? Were risks of bias and other limitations described?*

Implementation and Reporting: *Has the protocol provided a schedule for initial implementation? Has a schedule been provided for interim and final reporting of the claims evaluation?*

GUIDELINES FOR FORMULARY EVALUATIONS

5. NEXT GENERATION SEQUENCING PLATFORM ASSESSMENT PROTOCOL

5.1 Comparative NGS Platform Assessment

Any submission to a formulary committee that involves the identification of molecule markers and mutations has to justify the test or assay that supports target patient identification. This applies equally to single molecule tests as well as to NGS platforms. As NGS platforms (and the generations after that) are effectively the standard for genetic assessment, the questions that a formulary committee should ask are framed in the context of these platforms. With the growing number of platforms, it is important that these questions are framed in comparative terms. The focus of the comparative assessment is on: (i) the type and distribution of mutations and mutation clusters identified for a specific tumor type and (ii) the drugs or compounds that are matched to these mutations. Care has to be taken that poorly structured and managed NGS platforms are eliminated.

5.2 Evidence Base for Comparative NGS Assessment

Title: Evidence base for Comparative Assessment of [] NGS Platform

Abstract: *Structured summary of comparative assessment objectives: (i) description of NGS platform and comparator NGS platforms; (ii) NGS platform approvals; (iii) performance metrics (iii) target population for the comparative assessment; (iv) assessment criteria; (v) mutation distributions for the NGS platform and comparators; (vi) recommended therapy pathways for NGS platform and comparators; and (vi) conclusions regarding comparative value of NGS platform.*

NGS Platform Description: *a complete description of NGS platform procedures, library preparation, sequencing technology, signal detection and performance metrics*

Comparator NGS Platform Descriptions: *a description for each comparator NGS platform of procedures, library preparation, sequencing technology, signal detection and performance metrics*

NGS Platform Approvals: *A summary of the approval given for the NGS platform detailing evidence for analytical and clinical validity, approval as a medical device, any restrictions on application by disease area and target population, approvals for test in treatment guidelines, requirements for monitoring platform performance and processes for curating and updating test*

Comparator NGS Platform Approvals: *A summary of the approval given for each comparator the NGS platform detailing evidence for analytical and clinical validity, approval as a medical device, any restrictions on application by disease area and target population, approvals for test in treatment guidelines, requirements for monitoring platform performance and processes for curating and updating test*

GUIDELINES FOR FORMULARY EVALUATIONS

Target Population: *the proposed target population(s) for the NGS platform and comparator platforms, stage of disease, timeframe for evaluation and feedback*

Mutation Profile: *mutation profile for the target population(s) for NGS platform and comparators, distribution of target population(s) by mutation load for NGS platform and comparators, profile of comorbidities*

Therapy Profile: *matched therapy profile for mutation profile for NGS platform and comparators, therapy profile for comorbidities for target population(s)*

Assessment of Comparative Mutation and Therapy Profiles: *correspondence of mutation and therapy profiles, gaps*

Clinical Utility: *Comparative assessment of therapy choice outcomes by therapy pathway for proposed NGS platform versus comparator platforms in target population within a disease state*

Modeled Comparison: *where appropriate a common model should be applied to compare proposed NGS platform to comparator platforms to provide framework for claims and claims assessment to establish clinical benefits, resource utilization and costs in the target population*

Clinical benefits: *evaluable benefits defined for each therapy pathway in target population*

Resource Utilization: *claimed resource utilization profile for each therapy pathway for the proposed NGS platform and comparators*

Costs: *direct medical costs for the NGS platform and comparators defined for a fixed time interval, for each therapy pathway and overall*

Budget Impact: *projected budget impact of introducing NGS platform versus comparators in the target population and contribution to overall costs of treatment in the disease state*

Implementation and Claims Assessment: *reporting period for claims assessment*

5.3 Questions a Formulary Committee Should Ask

On receipt of an NGS comparative platform assessment proposal the formulary committee should consider:

Objectives: *Has the assessment provided a summary of the study objectives? Does the protocol abstract provide a summary of (i) population targeted; (ii) comparators; (iii) timeframe; (iv) study design; (v) data sources; (vi) endpoints; and (vii) reporting format?*

Context: *Has the assessment provided a statement for the need for the study and its benefits in the target population(s)?*

GUIDELINES FOR FORMULARY EVALUATIONS

Target Population: *Has the protocol provided a rationale for the population to be targeted, to include (i) demographic characteristics; (ii) clinical status; (iii) disease stage; (iv) mutation distribution; and (v) co-morbidities?*

Claims: *Has the protocol provided a summary of the modeled claims for outcomes to be evaluated? Has the protocol identified claims for (i) clinical benefits; (ii) adverse events and safety; (iii) resource utilization; (iv) direct medical costs; (v) cost-effectiveness and (vi) adherence and persistence for the therapy pathways in the target population?*

Study Design: *Has the protocol provided a rationale for the study design (e.g., prospective basket effectiveness trial, cohort/observational study)? Were alternative study designs proposed? Were risks of bias and other limitations described?*

Implementation and Reporting: *Has the protocol provided a schedule for initial implementation? Has a schedule been provided for interim and final reporting of the claims evaluation?*

6. REQUEST FOR SUBMISSIONS

These Guidelines are not intended to support unsolicited submissions and claims for clinical or cost-effectiveness. If a formulary committee wishes to evaluate a new product or NGS platform, or to revisit a product or NGS platform already on formulary, then it is at the discretion of the committee as to whether or not a request would be made. In the latter case submissions may be requested from all manufacturers of the comparator products and NGS platforms selected.

GUIDELINES FOR FORMULARY EVALUATIONS

REFERENCES

- ¹ Langley PC. Nullius in verba: The University of Minnesota Social and Administrative Pharmacy Program proposed Guidelines for Formulary Evaluations. *Inov Pharm.* 2016; 7(2): Article 10
- ² Wootton D. *The Invention of Science*. New York: Harper, 2015
- ³ Langley P. The status of modeled claims. *J Med Econ.* 2015;18(12):991-992
- ⁴ Langley P. Validation of modeled pharmacoeconomic claims in formulary submissions. *J Med Econ.* 2015;18(12):993-999
- ⁵ Academy of Managed Care Pharmacy. *Format for Formulary Submissions (Version 4)*. AMCP: April 2016
- ⁶ P Langley. Modeling imaginary worlds: Version 4 of the AMCP Format for Formulary Submissions. *Inov Pharm.* 2016;7(2): Article 11
- ⁷ National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE, April 2013
- ⁸ Langley PC. Sunlit uplands: The genius of the NICE reference case. *InovPharm.* 2016;7(2):Article 12
- ⁹ New Zealand. Pharmaceutical Management Agency. Prescription for Pharmacoeconomic Analysis: Methods for cost-utility analysis (v 2.1). PHARMAC, 2012.
- ¹⁰ Langley PC. He ao pohewa: The PHARMAC Prescription for Pharmacoeconomic Analysis in New Zealand and the standards of normal science. *Inov Pharm.* 2016; 7(2): Article 13
- ¹¹ Ireland. Health Information and Quality Authority. Guidelines for the Economic Evaluation of Health Technologies in Ireland. Dublin: HIQA; 2014
- ¹² Langley PC. Na domhain shamhlaíochta: formulary submission guidelines in Ireland and the standards of normal science. *Curr Med Res Opin.* 2016;32(5): DOI:10.1080/03007995.2016.1190699
- ¹³ Langley PC. Imaginary Worlds: The European Network for Health Technology Assessment (EUnetHTA) Recommendations for Health Economic Evaluations. *Inov Pharm.* 2016;7(3): Article 3.
- ¹⁴ Langley PC. Imaginary Worlds: Modeled Claims for cost-effectiveness published in *Value in Health* January 2015 to December 2015. *Inov Pharm;* 7(2): Article 18
- ¹⁵ Langley PC. Imaginary Worlds: Modeled claims for cost effectiveness published in *PharmacoEconomics* January 2015 to December 2015. *Inov Pharm.* 2016;7(2): Article 9
- ¹⁶ Langley PC. Imaginary Worlds: Modeled claims for cost-effectiveness published in the *Journal of Medical Economics* January 2015 to December 2015. *Inov Pharm.* 2016; 7(2):16

GUIDELINES FOR FORMULARY EVALUATIONS

- ¹⁷ Langley PC. Great Expectations: Cost-utility models as decision criteria. *Inov Pharm.* 2016; 7(2):14
- ¹⁸ Grim P, Rosenberger R, Eason RE et al. How simulations fail. *Synthese.* 2013;190(12):236-2390
- ¹⁹ Langley PC. Supporting Formulary Decisions: The discovery of new facts or constructed evidence? *INOVATIONS in Pharmacy.* 2016;7(2):15 <http://pubs.lib.umn.edu/Inovations/vol7/iss2/15>
- ²⁰ Popper KR. *Objective Knowledge* (rev. ed.). Oxford: Clarendon Press, 1959
- ²¹ Langley PC. Supporting formulary decisions: The discovery of new facts or constructed evidence? *Inov Pharm.* 2016;7(2): Article 15.
- ²² Centre for Evidence Based Medicine. *COMPare Project.* Oxford, 2016.
- ²³ Editorial. *Nature.* 2016;529:256
- ²⁴ Camerer CF, Dreber A, Forswell E et al. Evaluating replicability of laboratory experiments in economics. *Science.* 10.1126/science.aaf0918 (2016)
- ²⁵ Langley PC. Creative Destruction: Next Generation Sequencing in Drug Development, Formulary Evaluations and Pricing. *Inov Pharm.* 2016;7(4): Article 12
- ²⁶ Langley PC. Next Generation Sequencing Claims in precision medicine: Questions a formulary committee should ask. *Inov Pharm.* 2016;7(4): Article 11
- ²⁷ National Comprehensive Cancer Network (NCCN) <www.nccn.org>
- ²⁸ Pettitt D, Smith J, Meadows N et al. Regulatory barriers to the advancement of precision medicine. *Expert Rev Precision Med and Drug Development.* 2016;1(3):319-329
- ²⁹ Doran E, Hogue C. Potency, hubris and susceptibility; the disease mongering critique of pharmaceutical marketing. *Qual Report.* 2014;19(78): 1-18
- ³⁰ Howard DR, Bach PB, Berndt ER et al. Pricing in the market for anticancer drugs. *J Econ Perspectives.* 2015;29(1):139-162
- ³¹ Schnipper LE, Davidson NE, Wollins DS et al. Updating the American Society of Clinical Oncology Value Framework: Revisions and reflections in response to comments received. *J Clin Oncol.* 2016: DOI:10.1200/JCO.2016.68.2518
- ³² Bossuyt PM, Reitsma JB, Bruns DE et al. STARD 2015 – An updated list of essential items for reporting diagnostic accuracy studies. *BMJ.* 2015;351:h5527. PMID: [26511519](https://pubmed.ncbi.nlm.nih.gov/26511519/); Radiology. 2015;151:1516. PMID: [26509226](https://pubmed.ncbi.nlm.nih.gov/26509226/) Clinical Chemistry. 2015. pii: clinchem.2015.246280. PMID: [26510957](https://pubmed.ncbi.nlm.nih.gov/26510957/). See also [STARD item list](http://www.equator-network.org/reporting-guidelines/stard/) [<http://www.equator-network.org/reporting-guidelines/stard/>]
- ³³ Whiting PF, Rutjes A, Westwood M et al. QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Ann Int Med.* 2011;155:529-36. See also [<http://www.bristol.ac.uk/social-community-medicine/projects/quadas/>]

GUIDELINES FOR FORMULARY EVALUATIONS

- ³⁴ Moher D, Shamseer L, Clarke E et al. Preferred reporting items for systematic reviews and meta-analyses protocol (PRISMA-P) 2015 statement. *Systematic Rev.* 2015;4:1 DOI: 101186/2046-4053-4-1
- ³⁵ Yu W, Gwinn M, Dotson WD et al. A knowledge base for tracking the impact of genomics on population health. *Genet Med.* 2016, June 9 DOI: 10.1038/gim.2016.63. See also [<https://phgkb.cdc.gov/GAPKB/phgHome.do?action=home>]
- ³⁶ Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomized trials. *Ann Int Med.* 2010;152(11). See also: The CONSORT Statement 25-item check list [<http://www.consort-statement.org/checklists/view/32-consort/66-title>] and flow diagram [<http://www.consort-statement.org/consort-statement/flow-diagram>] to record the progress of patients through the trial.
- ³⁷ Meader N, King K, Llewellyn A et al. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. *Systematic Rev.* 2014;3:82
- ³⁸ Cochrane Handbook: [http://handbook.cochrane.org/part_2_general_methods_for_cochrane_reviews.htm].
- ³⁹ Berkman MD, Lohr KN, Ansari M et al. Grading the strength of a body of evidence when assessing health care interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An update. Methods Guide for Comparative Effectiveness Reviews (Prepared by the RTI-UNC Evidence-based Practice Center under Contract No. 290-2007-10056-I). AHRQ Publication No. 13 (14)-EHC130EF. Rockville MD: Agency for Healthcare Research and Quality. November 2013.
- ⁴⁰ Husereau D, Drummond M, Petrou S et al. Consolidated health economics evaluation reporting standards (CHEERS) statement. *J Med Econ.* 2013;16(6):713-19
- ⁴¹ Schommer J, Carlson A, Rhee G. Validating pharmaceutical product claims: questions a formulary committee should ask. *J Med Econ* 2015;18(12):1000-1006