Frequency and level of evidence used in recommendations by the National Comprehensive Cancer Network guidelines beyond approvals of the US Food and Drug Administration: retrospective observational study

Jeffrey Wagner,1 John Marquart,1 Julia Ruby,1 Austin Lammers,2 Sham Mailankody,3 Victoria Kaestner,4 Vinay Prasad2,4,5

OBJECTIVE
To determine the differences between recommendations by the National Comprehensive Cancer Network (NCNN) guidelines and Food and Drug Administration approvals of anticancer drugs, and the evidence cited by the NCCN to justify recommendations where differences exist.

DESIGN
Retrospective observational study.

SETTING
National Comprehensive Cancer Network and FDA.

PARTICIPANTS
47 new molecular entities approved by the FDA between 2011 and 2015.

MAIN OUTCOME MEASURES
Comparison of all FDA approved indications (new and supplemental) with all NCCN recommendations as of 25 March 2016. When the NCCN made recommendations beyond the FDA’s approvals, the recommendation was classified and the cited evidence noted.

RESULTS
47 drugs initially approved by the FDA between 2011 and 2015 for adult hematologic or solid cancers were examined. These 47 drugs were authorized for 69 FDA approved indications, whereas the NCCN recommended these drugs for 113 indications, of which 69 (62%) overlapped with the 69 FDA approved indications and 44 (39%) were additional recommendations. The average number of recommendations beyond the FDA approved indications was 0.92. 23% (n=10) of the additional recommendations were based on evidence from randomized controlled trials, and 16% (n=7) were based on evidence from phase III studies. During 21 months of follow-up, the FDA granted approval to 14% (n=6) of the additional recommendations.

CONCLUSION
The NCCN frequently recommends beyond the FDA approved indications even for newer, branded drugs. The strength of the evidence cited by the NCCN supporting such recommendations is weak. Our findings raise concern that the NCCN justifies the coverage of costly, toxic cancer drugs based on weak evidence.

Introduction
In the United States, the Food and Drug Administration may grant anticancer drugs or biologics either new drug marketing authorization—for compounds or agents not previously on the US market, or supplemental marketing authorization—approving a new indication for a drug already on the market. Beyond the FDA, modern oncologic care often involves the use of drugs or combinations that have not been explicitly approved—referred to as “off-label” use. Many off-label uses emerge outside of the normal drug development process and thus are not registered despite use in clinical practice. Off-label use accounts for sizable annual expenditures in the United States, and some estimate it is as much as half of all oncologic care. In an analysis of the 10 most commonly used cancer drugs, 30% of use was off-label, accounting for $4.5bn (£3.2bn; €3.6bn) in spending. Among off-label use, 46% (14/30), accounting for $2bn, was supported by guidelines from the National Comprehensive Cancer Network (NCCN). The NCCN is a prominent set of cancer specific guidelines used in clinical practice and now serves as one of five compendiums for private insurer and the Centers for Medicare and Medicaid Services (CMS) coverage. The process by which the NCCN issues its guidelines involves 54 individual panels comprising more than 1275 clinicians and oncology researchers from 27 member institutions. Per the NCCN website “The development of the NCCN Guidelines is an ongoing and iterative process, which is based on a critical review of the best available evidence and derivation of recommendations by a multidisciplinary panel of experts in the field of...
cancer.” The guidelines developed by these disease specific panel members result in algorithmic pathways delineating the process from staging through treatment and surveillance. The guidelines are updated at a minimum annually but often more frequently. For transparency, the NCCN posts short summaries of the meetings and panel discussions as well as disclosures of conflicts of interest. There is no mention of how the evidence is gathered or reviewed for recommendations.

In 1993, as part of the Omnibus Budget Reconciliation act, Congress mandated that CMS use three expert compendiums to determine coverage decisions for off-label drugs used in cancer care. In other words, CMS would reimburse for off-label use if the compendiums sanctioned that use. Over time, the number of compendiums has grown. As of 2016, there are now five: Wolters Kluwer Clinical Drug Information Lexi-Drugs, the American Hospital Formulary Service, Clinical Pharmacology, Micromedex DrugDex, and the National Comprehensive Cancer Network Drugs and Biologies accepted by the CMS. These compendiums are often used by commercial insurance providers for coverage decisions.

Multiple studies have evaluated the impact of such guidelines on clinical practice, supporting their use. In 2009, a systematic review of the compendiums found that the quality of evidence supporting recommendations was low. One study looked at a sample of 14 off-label recommendations and found they were supported by just one phase III study, 42 phase I or II studies, and three case reports. In 2010, a broader review of the strength of evidence of NCCN recommendations found that in general just 6% of recommendations were level 1—meaning they were of a high level with uniform consensus.

We examined the NCCN guidelines because of their widespread use in clinical practice. Specifically, for five consecutive years of FDA approved cancer drugs we determined what percentage of NCCN recommended uses fell within and outside of FDA guidance; the nature of the recommendations made when the NCCN recommended drugs beyond their FDA labels; the cited evidence in support of those recommendations; whether NCCN recommendations become more common over time; and the number of additional recommendations that lead to FDA approval. Given that NCCN recommendations support reimbursement for both commercial and CMS insurers, our investigation has direct health policy relevance.

Methods

Drug selection
We selected all new molecular entities approved for marketing by the FDA for the treatment of adult hematologic or solid cancers between 1 January 2011 and 31 December 2015. These years were chosen as they were the last five complete years at the time of our study, and drug approvals were obtained from FDA website (www.fda.gov/drugs/informationondrugs/approveddrugs/ucm279174.htm). We excluded changes in formulation, conversions of accelerated to regular approval, secondary approvals, and pediatric labeling.

FDA approvals
For each drug included in our dataset, we downloaded the most recent drug label as of 25 March 2016 and extracted all FDA approved new drug applications or biologic license applications. We included both original indications and supplementary indications up to 25 March 2016, and considered all of these to be FDA approved indications.

NCCN recommendations
We downloaded a pdf copy of the NCCN guidelines for the treatment of cancer by site for all available cancer types on 25 March 2016 (www.nccn.org/professionals/physician_gls/f_guidelines.asp). The length of all the NCCN guidelines in pdf format on this date was 3991 pages.

Comparing FDA approvals with NCCN recommendations
For each drug included in our dataset, we compared FDA approved indications as of 25 March 2016 against NCCN recommended indications as of 25 March 2016. We identified NCCN indications by searching the NCCN documents (3991 pages) for the drug’s generic name and examining each instance where the drug was mentioned. These instances were reviewed by a group, including two hematologist-oncologists. We identified the number of indications that overlapped (ie, where the NCCN recommendation and the FDA approval were identical), and the number of indications where the NCCN made additional recommendations beyond the FDA. We coded NCCN recommendations into one of five categories:

- Does not require prior treatment—that is, the NCCN recommendation did not mandate or specify some prior treatment, which the FDA did specify.
- Does not require concurrent treatment—that is, the NCCN recommendation removed the mandate to combine the agent with another drug specified by the FDA.
- Broadening the indication—that is, the NCCN recommendation removes other required molecular or clinical prerequisites, or both.
- Permits novel combination—that is, the NCCN recommendation allows physicians to combine the drug with another anticancer drug in a way not specifically endorsed by the FDA.
- Expands to a different malignancy—that is, the NCCN recommendation permits the use of the drug for a cancer other than that specified by the FDA.

For each recommendation, we then searched the NCCN guidance document for a reference supporting its use. We coded the evidence in support of these recommendations as one of: no evidence provided; book chapter or review article; case report or series of fewer than five people; case series of five or more people; phase I trial; phase II trial without randomization and fewer than 50 people; phase II trial.
without randomization and 50 or more people; phase II trial with randomization and fewer than 50 people; phase II trial with randomization and 50 or more people; randomized, phase III trial; and an ongoing clinical trial.

Follow-up of additional recommendations
On 19 December 2017, approximately 21 months from our original analysis, we examined all FDA drug labels for drugs for which the NCCN made additional recommendations. We identified how many additional recommendations led to FDA approval, and the evidence used by the FDA to grant those approvals.

Statistical analysis
Descriptive statistics are reported throughout. Statistical analysis was performed using STATA V.13.0 (StataCorp, TX). Our study was conducted between 25 March 2016 and 22 May 2017. Follow-up was performed in December 2017.

Patient involvement
No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study. No patients were asked to advise on interpretation or writing up of results. We plan to disseminate our results via email to some patient advocacy groups.

Results
We examined 47 drugs (new molecular entities) initially approved by the FDA between 1 January 2011 and 31 December 2015 for hematologic or solid cancers in adults. As of 25 March 2016 these 47 drugs were approved for a total of 69 FDA approved indications. The average number of FDA approved indications for each drug was 1.47. Thirty two (68%) of the drugs had one additional recommendation and 15 (32%) had two or more FDA approvals.

The NCCN recommended these drugs for 113 indications, of which 69 (61%) overlapped with the 69 FDA approved indications and 44 (39%) were additional recommendations. The average number of NCCN recommendations for these drugs was 2.40, and the average number of recommendations beyond the FDA approved indications was 0.92. Twenty one of the 47 (45%) drugs had no additional recommendations and 26 (55%) had one or more additional recommendations. Overall, 14 drugs (30%) had one additional recommendation, 6 (13%) had two additional recommendations, and 6 (13%) had three or more additional recommendations. The mean number of additional recommendations each year was 0.69, 0.89, 0.88, 0.91, and 1.67 for 2011, 2012, 2013, 2014, and 2015, respectively.

The NCCN made 44 recommendations beyond FDA approvals (see supplementary file for details). Among these recommendations, 13 (29%) were removing required prior treatment, 7 (16%) were removing required concurrent treatment, 8 (18%) were removing other inclusion criteria, 2 (4%) were permitting a novel combination, and 14 (32%) were for the treatment of a different malignancy. The 14 expansions to treatment of a different malignancy were made for 12 drugs (table 1). Eleven (92%) of these agents were targeted treatments and eight (75%) were known for specific enzymatic inhibition.

We sought to assess the cited evidence for the recommendation for these drugs. Only 23% (n=10) additional recommendations were cited as being based on evidence from randomized controlled trials, and just 16% (n=7) based on evidence from phase III studies (table 2).

Twenty one months after our analysis, we examined FDA labels to see how many additional recommendations from March 2016 had received FDA approval and the evidence base supporting those approvals. Six of 44 (14%) additional recommendations had received FDA approval. Four out of the six (66%)

<table>
<thead>
<tr>
<th>Drugs</th>
<th>FDA cancer approval</th>
<th>NCCN cancer recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palbociclib</td>
<td>Hormone receptor (HR)-positive, human epidermal growth factor receptor 2 (HER2)-negative advanced breast cancer</td>
<td>Soft tissue sarcoma</td>
</tr>
<tr>
<td>Centinib</td>
<td>Anaplastic lymphoma kinase (ALK)-positive metastatic non-small cell lung cancer (NSCLC)</td>
<td>Soft tissue sarcoma</td>
</tr>
<tr>
<td>Trametinib</td>
<td>Melanoma with BRAF V600E or V600K mutations as detected by an FDA approved test</td>
<td>NSCLC</td>
</tr>
<tr>
<td>Dabrafenib</td>
<td>Metastatic melanoma with BRAF V600E or V600K mutation</td>
<td>NSCLC</td>
</tr>
<tr>
<td>Radium-223</td>
<td>Castration resistant prostate cancer, symptomatic bone metastases, and no known visceral metastatic disease</td>
<td>Osteosarcoma</td>
</tr>
<tr>
<td>Pomalidomide</td>
<td>Multiple myeloma</td>
<td>Systemic light chain amyloidosis</td>
</tr>
<tr>
<td>Cabozantinib</td>
<td>Progressive, metastatic medullary thyroid cancer</td>
<td>Kidney cancer; NSCLC</td>
</tr>
<tr>
<td>Carfilzomib</td>
<td>Relapsed or refractory multiple myeloma</td>
<td>Waldenstrom's macroglobulinemia; lymphoplasmacytic lymphoma</td>
</tr>
<tr>
<td>Crizotinib</td>
<td>Locally advanced or metastatic NSCLC that is anaplastic lymphoma kinase (ALK) positive, metastatic NSCLC with tumors that are ROS1 positive</td>
<td>Soft tissue sarcoma</td>
</tr>
<tr>
<td>Brentuximab vedotin</td>
<td>Hodgkin lymphoma; anaplastic large cell lymphoma; classic Hodgkin lymphoma</td>
<td>Mycosis fungoides/Sezary syndrome</td>
</tr>
<tr>
<td>Venetoclax</td>
<td>Unresectable or metastatic melanoma with BRAFV600E mutation</td>
<td>Hairy cell leukemia; NSCLC</td>
</tr>
<tr>
<td>Vandetanib</td>
<td>Symptomatic or progressive medullary thyroid cancer in patients with unresectable locally advanced or metastatic disease</td>
<td>Papillary or Hurthle cell thyroid cancer (non-medullary)</td>
</tr>
</tbody>
</table>
did so on the basis of data from randomized controlled trials, and two did so (33%) on the basis of data from phase II trials with more than 50 participants.

**Discussion**

On average, among cancer drugs approved over the preceding five years, the NCCN endorses approximately one recommendation beyond the median of 1.5 FDA granted approvals. When the NCCN makes recommendations beyond the FDA approval, it most commonly does so by removing a prior treatment, removing inclusion criteria, or expanding treatment to a different malignancy.

These recommendations are often based on low quality of evidence or no evidence. Just 23% (10/44) of these recommendations are supported by evidence from randomized controlled trials, while most do not provide references or are based on small, uncontrolled studies or case reports. In contrast, among 83 consecutive cancer drug approvals by the FDA over a similar period, 58% (n=48) were based on evidence from randomized controlled trials. Thus, we found that citations provided by NCCN for additional recommendations beyond the FDA approval appeared to be based on less robust trials than approvals by the FDA. Given that NCCN endorsement is linked to reimbursement by many commercial insurers and the CMS, our results suggest that payers may be covering cancer drugs with varying and scientifically less robust justification. Some of the recommendations beyond FDA approval expand the population of patients treated, but many of these patients may be better served by participation in controlled trials. With extended follow-up of 21 months, few additional recommendations of the NCCN led to FDA approvals (6/44, 14%).

Low quality of evidence supporting oncology practices has led to several noted medical reversals (where a widespread intervention was found to be no better than previous care) as well as inflated estimates of effect sizes. Specifically, the use of high dose chemotherapy and salvage autologous stem cell transplant for women with breast cancer gained prominence by uncontrolled, phase II studies, before being contradicted by at least six randomized controlled trials. Evidence also shows that a measure of drug activity, the response rate, is consistently larger in phase II trials compared with phase III studies. Thus, the NCCN’s reliance on lower quality of evidence may lead to false inferences concerning the efficacy of toxic, costly cancer drugs or the magnitude of their benefit. Although it is tempting to conclude that extrapolating a drugs benefit to a different malignancy (based on biologic rational) is more egregious than extrapolating a drugs benefit in a previous line of treatment, the reality is both are empirical questions that involve recommending a drug for which cost and toxicity is certain, but benefit unknown.

The use of compendiums to justify coverage was once rational. In the early 1990s, nearly all anticancer drugs were relatively low cost, cytotoxic treatments, which were used in diverse malignancies based on some evidence. Since many of these drugs were older or off-patent, the impetus to locate funding to seeking formal FDA approval was low. However, the same rationale may not apply equally to newer drugs. Nearly all recommendations in our dataset are related to targeted therapy and call into question whether off-label use based on a potential target is appropriate. These drugs are branded, and typically cost in excess of $100,000 per year of treatment. Thus broadening of the indications by the NCCN may decrease the incentive to carry out further studies that would clarify the evidence. This may contribute to the lower quality of evidence in support of additional recommendations. For a new indication, we feel it is justified to advocate for an increased number of trials to determine the appropriateness in the clinical setting.

Finally, recent research has drawn attention to the role of financial conflicts of interest in the development of compendiums. Specifically, 86% of NCCN guidelines members have financial ties to the industry, with 84% receiving personal payments and 47% receiving research payments. The presence of conflicted physicians has been shown to lead to more optimistic conclusions regarding disputed practices. Thus our findings raise concern about the nature of the recommendations offered by these individuals.

**Strengths and limitations of this study**

We reviewed all NCCN recommendations compared with FDA approvals. We additionally evaluated the cited evidence used to support recommendations made by the NCCN, and reanalyzed FDA approvals 21 months after our initial study.

Our study has four key limitations. First, we did not assess which recommendations were NCCN level 1, 2A, 2B, or 3, as these were not always clearly listed in the guidelines, and the level may have implications for coverage decisions. For instance, different insurers have different thresholds for coverage. United Healthcare and Aetna cover all recommendations higher than 2B. Cigna covers all recommendations 2A or greater and decides 2B recommendations on a case by case basis. The CMS states only level 3 recommendations are “not medically accepted” and
not covered.\textsuperscript{20} However, others have shown that 89% of NCCN recommendations are level 1 or 2A, 99% are 2B or better, and just 1% of recommendations are level 3.\textsuperscript{7} Thus, even had we been able to extract the strength of the recommendation, most of these recommendations are likely to involve coverage.

Second, we did not search for independent evidence to support recommendations beyond the references provided by the NCCN. Given that the NCCN is the entity making recommendations and as we are uncertain beyond listed references what studies were considered in the decision making process, we felt that it was the obligation of the guideline authors to provide the best evidence in support of their recommendations. Additionally, the task for our research team of performing 44 systematic reviews was considered a prohibitive time commitment in characterizing the literature related to each recommendation. Notably, 21 months later just six of the 44 additional recommendations were included in updated FDA drug labels.

Third, our analysis was performed on a specific date and time, locking FDA approvals and NCCN recommendations. In the 21 months that followed, some of the NCCN additional recommendations received FDA approval (6/44). Others may have had additional data generated, but this was beyond the scope of our investigation. Yet, many recommendations may continue to lack empirical support. When it comes to costly cancer drugs with serious toxicities, it is important to recognize that merely because some drugs were later validated does not mean that off-label recommendations for all drugs is, on average, a good thing. It remains an open question.

Lastly, we recognize our findings may not only cause oncologists to question the underlying evidence for current guidelines but also will do the same for patients. The complex considerations oncologists and patients make in relation to goals of care and expected clinical outcome takes into account a wide variety of variables. To help patients navigate the discussion of whether a potential treatment plan is supported by quality evidence, we recommend they address these concerns directly with their providers.

Conclusion and policy implications

We systematically compared drug approvals by the FDA with recommendations by the NCCN, and further characterized which of those recommendations went beyond the original FDA approval. We further attempt to determine the cited evidence for which the NCCN used to make its recommendation based on those listed in the guidelines, analyzing what recommendations went beyond the FDA approval and also had low quality evidence cited or absent. With longer follow-up, the number of NCCN low quality evidence recommended recommendations may grow.

When the NCCN makes such recommendations, it does so often by removing inclusion criteria or prior required treatment, or expands treatment to a different malignancy. These actions increase cost and toxicity but have an unknown impact on outcomes. The cited evidence to support these recommendations is generally poor, with heavy reliance on uncontrolled studies, case reports, expert opinion, or no offered evidence. If there is additional evidence in support of these recommendations the NCCN should improve its process and cite all evidence used. Given the important role of the NCCN in guiding coverage decision both for commercial insurers and for the CMS, we believe the standards for compendiums inclusion warrant independent audit.

Contributors: VP designed the study, analyzed data, ensured accuracy of analysis, and drafted the manuscript. The primary authors (JW, JM, JR, and AL), who contributed equally to the study, initiated the collaboration, assisted in developing the research concept, completed guideline review, performed data collection plus analysis, and drafted the manuscript in collaboration with the primary, expert, build upon UK and SM provided comments, revised the research concept, and revised the manuscript. JW, JM, JR, AL, and VP are the guarantors.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. VP is funded by the Laura and John Arnold Foundation for work on medical reversal.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: Data used for analysis are in the supplementary table.

Transparency: The manuscript’s guarantors (JW, JM, JR, AL, and VP) affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

9 Carroll J. Biologics debate heats up between FDA, NCCN. Manag Care 2011;20:30-2.
15 Mailankody S, Prasad V. Five years of cancer drug approvals: Innovation, efficacy, and costs. JAMA Oncol 2015;1:539-40. 10.1001/jamaoncol.2015.0373
18 Fugh-Berman A, McDonald CP, Bell AM, Bethards EC, Scialli AR. Promotional tone in reviews of menopausal hormone therapy after the Women’s Health Initiative: an analysis of published articles. PloS Med 2011;8:e1000425. 10.1371/journal.pmed.1000425
19 De Jesus-Morales K, Prasad V. Closed Financial Loops: When They Happen in Government, They’re Called Corruption; in Medicine, They’re Just a Footnote. Hastings Cent Rep 2017;47:9-14. 10.1002/hast.700