Determining the efficacy of dietary phytochemicals in cancer prevention

Margaret M. Manson*, Bethany E. Foreman, Lynne M. Howells, Elena P. Moiseeva
Cancer Biomarkers and Prevention Group, Department of Cancer Studies and Molecular Medicine, Biocentre, University of Leicester, LE1 7RH.

*corresponding author

mmm2@le.ac.uk

Keywords: chemoprevention, biomarkers, indoles, polyphenols, tumour subtypes, plasma profiling

Abbreviations: DIM – di-indolylmethane; EGCG – epigallocatechin-3-gallate; EGF – epidermal growth factor; EGFR – epidermal growth factor receptor; EMT - epithelial to mesenchymal transition; ER – oestrogen receptor; ESA – epithelial specific antigen; Her2 – EGFR family member ErbB2; I3C – indole-3-carbinol; MS/MS - tandem mass spectrometry; SVM – support vector machine; uPA – urokinase plasminogen activator; VEGF – vascular endothelial growth factor
Abstract
Accumulating data suggest that dietary phytochemicals have the potential to moderate
deregulated signalling or reinstate checkpoint pathways and apoptosis in damaged cells,
while having minimal impact on healthy cells. These are ideal characteristics for
chemopreventive and combination anti-cancer strategies, warranting substantial research
effort into harnessing the biological activities of these agents in disease prevention and
treatment. However, this requires further investigation into their mode of action and novel
approaches to the development of reliable biomarkers.

Introduction
Epidemiology indicates that the incidence of many common cancers is geographically and
environmentally determined and that diets rich in fresh fruit and vegetables confer a lower
risk of developing tumours in many target tissues. Around one third of all cancers are
thought to be related to unhealthy diet and therefore, in theory, preventable [1]. However,
linking particular dietary components to prevention of specific cancers has proved to be a
major challenge.

Many phytochemicals of differing chemical structure have anticarcinogenic properties.
They activate cytoprotective enzymes and inhibit DNA damage to block initiation in
healthy cells, or modulate cell signalling to eliminate unhealthy cells at later stages in the
carcinogenic process. In vitro data for several well-studied compounds indicate that each
can affect many aspects of cell biochemistry, but different agents have many similar
activities (Table 1). Nevertheless effects can be cell-type specific, with transformed cells
being particularly sensitive. But despite the exponential increase in knowledge regarding
mechanisms of action of these molecules, their success in clinical trials has been limited.
In addition to the difficulties of long-term prevention trials in healthy populations, there are
other reasons for this. In cancer therapeutics, emphasis is increasingly placed on targeting
tumour subtypes within a particular tissue, but this has not generally been considered for
chemopreventive trials. In many instances the crucial in vivo targets for particular dietary
molecules are unknown, making it difficult to predict which cancer phenotypes are most
likely to be affected. Many phytochemicals are poorly bioavailable and evidence suggests
that combinations may be more effective than single agents. There may also be advantages
in combining them with chemo- or radio-therapies. Probably the greatest hindrance to
more successful trials is the lack of validated biomarkers of efficacy. To enable many
more agents to be tested much more quickly, validated surrogate endpoint biomarkers are

2
required, which will accurately determine outcome at a much earlier time in the process of tumour development.

**Identifying at risk individuals**

In order to optimise the chances of success in cancer chemoprevention trials, the ability to identify those individuals most likely to benefit is clearly important [2]. In the case of primary prevention to inhibit the earliest stages of tumour development, selection has traditionally been based on known environmental and lifestyle risk factors, genetic predisposition and family history. But with such cohorts, obtaining definitive evidence to directly link exposure to a particular dietary agent with prevention of cancer in any target tissue is at best extremely challenging, at worst impossible. In future it is possible that individuals suitable for primary prevention will be identified through screening for more common multiple susceptibility loci, such as those recently reported for breast cancer [3]. Such loci together could inform on a substantial fraction of the genetic variance in some cancers.

Patients can be recruited at later stages of the carcinogenic process. Secondary prevention is appropriate for those who have already developed premalignant lesions, such as intraepithelial neoplasia or intestinal polyps, the progress of which can be monitored in response to chemopreventive treatments. Several dietary compounds, including indoles and polyphenols have shown promise in this respect, with regression of respiratory papillomatosis, cervical, vulvar and prostate intraepithelial neoplasia and oral leukoplakia [4-7]. However, the relationship of some of these early lesions to tumour outcome is uncertain. Clearly a strategy involving susceptibility loci would also be relevant at this stage, adding a degree of certainty to the prognosis.

A third strategy is tertiary prevention, which focuses on patients who have been successfully treated for a primary tumour, in order to inhibit development of second primary tumours. Greatest success to date in this respect has resulted from the use of drugs such as tamoxifen and its analogues for breast cancer, and retinoids for skin, head and neck and liver cancer [8]. If phytochemicals have a role at this stage, it is most likely to be as part of a combined therapy.

Arguably a fourth category of patients who could benefit would be those with a superficial or primary tumour, where intervention is designed to prevent invasion and metastases. While the cellular process of epithelial to mesenchymal transition (EMT) is fundamental to morphogenesis, when reactivated in cancer, it facilitates invasion and metastasis[9]. Reversing this process could limit metastatic spread, achieving late-stage prevention with
enhanced survival. Ability of chemopreventive agents to inhibit EMT or angiogenesis has been demonstrated in a number of animal models. Use of phytochemicals for this purpose would be facilitated by knowledge of the characteristics within the primary tumour which predict invasive potential. We previously showed that the polyphenol, curcumin, modifies the invasive potential of breast cancer cells[10]. Another polyphenol, epigallocatechin-3-gallate (EGCG), was found to inhibit neovascularisation in the chick chorioallantoic membrane assay and when given in drinking water could significantly suppress vascular endothelial growth factor (VEGF)-induced corneal neovascularisation [11]. Such results suggest that EGCG may be a useful inhibitor of angiogenesis in vivo. A number of phytochemicals also affect expression of cadherins, catenins and matrix metalloproteinases (Table 1), and prognostic metastatic biomarkers for breast cancer, such as urokinase plasminogen activator, uPA/PAI1 and Her2 [12] all of which can modulate invasive capacity.

**Tumour subtypes**

A key aspect of targeted therapies, which has so far received much less attention in chemoprevention trials, is the concept of tumour subtypes. Tumour development involves the accumulation of multiple mutations which differ from one tumour to the next in the same target tissue. Such changes can be assessed using gene expression profiles to identify patterns of pathway deregulation which inform on disease prognosis and indicate treatment options [13]. Subtypes are perhaps best defined for the leukaemias [reviewed in 14] and in breast cancer, where at least 5 subtypes including luminal A (oestrogen receptor α (ERα)-positive), luminal B (ERα-positive), basal-like (ERα-negative, Her2-negative, cytokeratin 5/6-positive, and/or epidermal growth factor receptor (EGFR)/Her1-positive), Her2 (Her2 amplicon-positive, ER-negative) and normal breast-like have been identified [15,16]. Such patterns in cancer cell lines have also been shown to be useful in predicting response to therapeutic agents [13, 17,18]. However, subtypes identified on the basis of gene expression profiles do not correspond exactly to those identified by immunohistochemistry [19,20] and used in clinical practice [21]. Additionally, gene signatures which predict the response to individual agents can be used in combination to predict the efficacy of multidrug regimes [22]. Dietary chemopreventive agents certainly exhibit cell-type specificity, but because of their broad ranging activity, they may be effective against different tumour subtypes. Recent studies show that indole-3-carbinol (I3C) induces apoptosis by decreasing expression of
genes essential for tumour cell viability, such as ERα and EGFR in breast cells of luminal A and basal-like subtypes [23,24]

**Molecular biomarkers**

There is no doubt that identification of reliable molecular markers, such as those in subtype molecular signatures, would be extremely advantageous to accurately predict efficacy of any intervention. Ideal markers should be indicative of early changes, relate directly to the carcinogenic process and where possible allow less invasive assessment of chemopreventive efficacy [25]. For some approaches, this requires a detailed knowledge of not only the stages of carcinogenesis for a particular tumour subtype, but also the mechanisms of action of the preventive agent. Molecular biomarkers would be valuable as targets to identify new agents or to optimise lead compounds; as risk biomarkers for selecting suitable cohorts for chemopreventive trials; or as indicators of efficacy for determining response to mechanism-based interventions or identifying potential toxicity.

There is no shortage of candidate proteins related to oncogenic processes (drug metabolising enzymes, growth factors, transcription factors, cell cycle and apoptosis related proteins) known to be aberrantly expressed in various tumours and modulated by phytochemicals *in vitro* (Table 1). The philosophy of chemoprevention suggests long-term or life-time exposure to low doses. Therefore, when selecting potential biomarkers of efficacy, it is essential to ensure that they are modulated under appropriate conditions. However, many of the *in vitro* mechanistic studies have been carried out with single high doses of chemopreventive agent that are not achievable *in vivo*. Thus some of the reported effects may not be physiologically relevant. In order to better predict how phytochemicals may act in humans, several approaches are required. Pharmacokinetic and pharmacodynamic studies allow estimates of plasma and, in some cases tissue, concentrations to be made. Where these already exist, they indicate that achievable concentrations are in the nanomolar, or at best, low micromolar range. In humans curcumin levels are typically in the low nanomolar range, although ~10 μM can be achieved in colorectal tissue. Resveratrol plasma concentrations of ~2μM appear typical [reviewed in 26]. These low doses then need to be applied in cell culture studies over extended periods to better mimic *in vivo* exposure. We exposed the metastatic breast line, MDA-MB231, to physiologically achievable concentrations of five agents in long-term culture and observed favourable alterations to cell cycle, clonogenicity, apoptosis and expression of several proteins associated with EMT (Moiseeva et al unpublished data). Encouragingly, quite a
few published studies, in which single treatments with physiological concentrations of dietary agents have been used *in vitro*, also reveal significant biological activity [reviewed in 26]. Additionally, a number of phytochemical effects observed *in vitro* have been validated in animal models.

Monolayer cultures are unrepresentative of *in vivo* environments, so that models where cells are grown on relevant substrata like collagen, laminin, or matrigel, or as 3-dimensional spheroids may be more compatible with *in vivo* activity. We found that breast cells responded differently to I3C under such growing conditions [27]. In 3D culture (collagen 1 or spheroids) the sensitivity of MCF7 and MDA-MB-468 breast cancer cells towards I3C was increased. In MDA-MB-468 cells the expression of the EGFR and β1-integrin was modulated by 3D culture, with cells responding differently to EGF or the EGFR inhibitor, PD153035. Phytochemical effects are often cell-type specific and so different panels of biomarkers may be required for different target tissues, or for different cancer subtypes within a single tissue. On the other hand, there appears to be a certain degree of similarity in the protein targets affected by a variety of structurally unrelated phytochemicals, suggesting similar mechanisms of action (Table 1).

A detailed understanding of the effects of dietary agents (for example on growth factor signalling, EMT, cell cycle arrest and apoptosis) following extended treatment at physiologically achievable doses, and related to target tissue and cancer subtype, will help to identify useful biomarkers. Such an understanding would include identification of primary targets of phytochemicals (particularly proteins such as receptors, or more general effects such as endoplasmic reticulum stress or altered redox status), an appreciation as to why healthy cells are generally more resistant and comparison of *in vitro* with *in vivo* efficacy.

**Cancer-initiating stem cells**

An increasing number of studies are reporting the identification of a subset of cancer stem cells within a tumour, which are thought to be responsible for the highly aggressive nature of different cancers. Breast cancer cells, grown in immune compromised mice, contained a subpopulation expressing cell surface markers [epithelial specific antigen (ESA)⁺CD44⁺CD24⁻/low Lineage⁻]. This subset maintained the ability to form new tumours [28 Al-Hajj]. Interestingly, the basal-like breast cancer subtype has similarity with breast stem cells [20, 29], suggesting the “stem-like” characteristics of this subtype may be responsible for its aggressiveness and poor prognosis. Using a xenograft model of primary human pancreatic adenocarcinoma, Li et al [30] described a subpopulation of cancer cells
expressing CD44, CD24 and ESA] that were 100-fold more tumorigenic than other cells from the same tumours. Similarly O’Brien et al., [31] identified a highly tumorigenic subset of colon cancer cells positive for CD133 (1 colon cancer initiating cell per $5.7 \times 10^4$ tumour cells). Clearly if biomarkers can be used reliably to identify subsets of cancer initiating cells, it will be essential to determine the efficacy of treatments against these, in preference to other phenotypes within the tumour.

**Dietary agents in combined treatments**

Also of increasing importance is the investigation of combinations of phytochemicals, or their use in conjunction with other therapies, to increase efficacy or decrease unwanted side effects [26, 32]. We have shown in breast cell lines that I3C exhibits enhanced efficacy in combination with src or EGFR kinase inhibitors [24] and *in vivo* I3C prevented the hepatotoxicity of trabectedin (ET743), an experimental anti-tumour drug with promising activity in sarcoma, breast and ovarian carcinomas, without compromising anti-tumour efficacy [33]. Curcumin enhances the efficacy of oxaliplatin in both p53<sup>positive</sup> and p53<sup>mutant</sup> colon cancer cells [34]. However, caution is required, since it has also been reported to compromise the efficacy of some chemotherapeutic drugs in human breast cancer models [35].

**Proteomics**

Methods which do not require a detailed knowledge of mechanisms, initially at least, offer an alternative approach to developing biomarkers and assessing chemopreventive efficacy. Mass spectrometry of biological samples offers a powerful proteomic tool for the discovery of novel biomarkers and for profiling [36]. Proteins and peptides in clinical samples reflect the intracellular activities of healthy and diseased tissue. Plasma samples subjected to mass spectrometry can provide characteristic protein profiles, when over- and under-expressed peptides are determined by pattern comparison using a variety of machine learning algorithms [37-39]. This technique should offer the possibility of monitoring plasma profiles in at risk individuals in response to a particular treatment regime. Development of these methods for cultured cells, would also provide the means to screen effects of chemopreventive agents – for example, to estimate the lowest concentration of agent that causes a change in profile; to look for dose-response in altered peaks; to compare effects of one agent in different cell types; to compare the signatures between compounds in the same cell type; and to enable identification of signature peaks. Individual proteins can be identified by correlating the sequences of tryptic peptides generated by tandem mass spectrometry (MS/MS) with sequences in protein databases.
The evolutionary algorithms with support vector machine (SVM) supervised learning enable identification of biomarkers with the highest diagnostic and prognostic potential. The effectiveness of combined treatments could also be explored. Nuclear fractions could provide enrichment for investigating changes in transcription factors and cell cycle-related proteins. In view of the importance of protein phosphorylation in proliferation, differentiation and apoptosis, development of techniques selective for phosphopeptides would provide added value [40, 41].

**Conclusions**

In order to fully appreciate the potential for dietary compounds in cancer prevention and bridge the gap between apparent *in vitro* efficacy and clinical use, a number of approaches are required. Attention must be paid to the use of physiologically relevant concentrations, to chronic exposure and to 3D cultures, to more accurately mimic *in vivo* situations. Validation of potential mechanisms in appropriate animal models is also important. The mechanistic biomarkers identified *in vitro* need to be verified in human tissues as being central to the carcinogenic process. Finally genomics and proteomics approaches offer novel ways of predicting clinical efficacy.

**Acknowledgements**

The authors are grateful to Professor Peter Farmer and Dr Don Jones for critical discussion.

**References**

Table 1 – Proteins modulated by dietary agents

<table>
<thead>
<tr>
<th>Dietary compound</th>
<th>Altered expression/activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>curcumin (from turmeric)</td>
<td>p21; p27; p38; p53; p70S6K; abl; AIF; pAkt; AP-1; ARNT; ATPase; Bad; Bax; Bel-2; Bel-X1; Bid; BTG2; E-cadherin; cadherin-11; caspase 3, 4, 8, 9; β-catenin; CD80/86; cdc2; cdc25c; cdc25b; CHK2; chymase II; collagenase; cyclin A/B1/D1/E; cty c; COX-2; DR5; EGFR; ElF2α; elastase; ELK1; Erg1; ERK1/2; FLIP; c-fos; GADD153/CHOP; GADD45; Gb3 synthase; GM-CSF; GRP78; HAT; HDAC; Her2; HIF-1; HO-1; hyaluronidase; IAP1/2; ICAM-1; IFNγ; IGF-1R; IkBα; IKK; IL2/6/8/10/12; iNOS; IRAk; IRF3; JAK1/2; JNK; c-jun; LOX; MAPKAP-K1β; MDM2; MHC; class II antigens; MMP2/3/9/13; mTOR; myc; MyD88; myeloperoxidase; Nag-1(PLAB); NF-κB; Notch1/3/4; Nr2; ODC; PARP; PCNA; PERK; PGE; phase1/2 enzymes; PHK; PKC; PPARγ; pRB; src; smac; STAT1/3; survivin; SYK; TGFB; Th1; thiorodoxin reductase; TNFa; TNFR1; topoisomerase II; TRAIL; α/β-tubulin; VEGF; wee1; XIAP;</td>
</tr>
<tr>
<td>resveratrol (from red grapes, wine)</td>
<td>p21; p27; p38; p53; p57; p70S6K; pAkt; AP-1; APAF1; ASK1; ATM/ATR; Bad; Bak; Bax; Bel-2; Bel-X1; Bid; caspase 2,3,8,9; E-cadherin; VE-cadherin; β-catenin; p300/CBP; cdc2; cdc25c; CDK2/4/6/7; CHK1/2; cyclin A/B1/D1/E; cty c; EGFR; ER; ErbB3; ERK; Fas; c-fos; Fasl2/2; GADD153/CHOP; GADD45a; GJC; GSK3; H2A.X; Her2; HO-1; IAP1; IkBα; IKK; IL6/8; iNOS; JNK; c-jun; LOX; Mad21; MDM2; MMP2/9; NF-κB; Nr2; ODC; PI3K; phase1/2 enzymes; PHK; PIG7; PIG8; PIG10; PKC; PPARγ; PR; PTEN; pRb; SIRT1; Sp1; src; STAT3; survivin; TOP2A; TRAIL R; TRIF; VEGF; XIAP;</td>
</tr>
<tr>
<td>EGCG (from green tea)</td>
<td>p16; p18; p21; p27; p38; p53; p130; p107; p70S6K; pAkt; AP-1; APAF1; Bad; Bax; Bel-2; Bel-X1; caspase 3,8,9; DNMT; E2F; E-cadherin; VE-cadherin; β-catenin; cdc2; CDK2/4/6/7; C/EBPα &amp; β; COX2; cyclin A/B1/D1/E; cty c; DP1/2; EGFR; Erg1; ERK1/2; Fas; Fasl1; βFGF; FGFR; fibrinogen; fibronectin; FKHR; c-fos; GADD153/CHOP; GADD45a; Her2; C/EBPα &amp; β; E-cadherin; caspase 3,8,9; cdk2; cdk25c; cdk25a; CDK2/4/6; cdc2; cdc25c; cdc25b; CHK2; chymase II; collagenase; cyclin A/B1/D1/E; cty c; COX-2; DR5; EGFR; ER; Ets; FLIP; Her2; IAP1/2; IL6/8; IκB; IKK; IL6/8; iNOS; involucrin; JNK; c-jun; lamnin; LOX; 67LR; Mcl-1; MDM2; MMP1; MMP2/7/9/13/14; NF-κB; Nr2; ODC; PDGFR; PGES-1; PI3K; phase1/2 enzymes; PKA; PKC; PRAK; RARβ; pRB; STAT3; survivin; telomerase; TIMP1; topoisomerase I; VEGF; VEGFR; vimentin; XIAP;</td>
</tr>
<tr>
<td>I3C (cruciferous vegetables)</td>
<td>p15; p16; p21; p27; p53; pAkt; ATM; Bad; Bax; Bel-2; Bel-X1; BCRP/ABCG2; BRCA1; E-cadherin; caspase 3,8,9; cdk2; CDK2/4/6; CHK2; COX2; CXCR4; cyclin D1/E; cty c; DR5; EGFR; ER; Ets; FLIP; Her2; IAP1/2; IL6/8; MDM2; MMP2/9; MUC1; Nag-1(PLAB); NF-κB; Nr2; ODC; Pgp; phase1/2 enzymes; PI3K; PTK; PTK; pRb; Sp1; src; STAT3; survivin; topoisomerase II; TRAF1; TRAIL R; uPA; VEGF; vimentin; XIAP;</td>
</tr>
<tr>
<td>DIM (cruciferous vegetables)</td>
<td>p21; p27; p38; pAkt; AR; ATF2; ATF3; Bax; Bel-2; Bel-X1; Bid; BRCA1; E-cadherin; cadherin-11; N-cadherin; P-cadherin; caspase 3,6,7,8,9; β-catenin; cdc2; cdc25a; CDK2/4/6; CREB; cyclin D1; cty c; DR5; EGFR; ER; ERK1/2; Ets; FLIP; GADD45; GADD153/CHOP; GRP78(Bip); Her2; JNK; c-jun; MEK; MHC-1; mitochondrial H+ -ATPase; MMP9; Nag-1(PLAB); NF-κB; p56/p69; oligoadenylate synthases; PCNA; phase1/2 enzymes; PI3K; PSA; Raf; Ras-GTP; smac/Diablo; Sp1/3; STAT3/5; survivin; TGF-α; topoisomerases IIα/IIβ; α/β-tubulin; uPA; VEGF;</td>
</tr>
</tbody>
</table>

The expression, phosphorylation, activity or binding of the proteins listed is affected by five of the best-studied dietary agents - data are taken from in vitro studies in many cell types, but the list is not comprehensive. There are many mechanistic similarities between compounds of different chemical structures indicated in bold. We apologise to all the authors for not citing the original references owing to space restrictions. DIM – di-indolylmethane.